

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:18 ; Search time 5179.29 Seconds  
(without alignments)  
9899.970 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctcgaagacc.....tctcttggatgccttag 1183

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_hg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_hg\_hum:\*
- 31: em\_hg\_inv:\*
- 32: em\_hg\_other:\*
- 33: em\_hg\_mus:\*
- 34: em\_hg\_pln:\*
- 35: em\_hg\_rtd:\*
- 36: em\_hg\_man:\*
- 37: em\_hg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_hgo\_hum:\*
- 40: em\_hgo\_mus:\*
- 41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1183	100.0	1183	6	AX443133	AX443133 Sequence
c	1183	100.0	1183	6	AX443135	AX443135 Sequence
3	1181.4	99.9	2567	6	AX452880	AX452880 Sequence
4	1152.4	97.4	2788	6	AX780857	AX780857 Sequence
5	1144.8	96.8	2538	9	BC042041	BC042041 Homo sapi
6	825.8	69.8	2415	9	AX025645	AX025645 Homo sapi
7	784.4	66.3	786	6	AX511153	AX511153 Sequence
8	784.4	66.3	786	6	AX572845	AX572845 Sequence
9	784.4	66.3	786	9	AF290985	AF290985 Homo sapi
10	784.4	66.3	786	9	AF326353	AF326353 Homo sapi
11	661.8	55.9	737	6	AX511155	AX511155 Sequence
12	661.8	55.9	737	9	AF290986	AF290986 Homo sapi
13	563	47.6	1384	10	BC052655	BC052655 Mus muscu
14	561.4	47.5	1348	6	AX511150	AX511150 Sequence
15	561.4	47.5	1348	10	AF287467	AF287467 Mus muscu
16	559.8	47.3	1321	10	AF434990	AF434990 Mus muscu
17	535.2	45.2	777	6	AX511151	AX511151 Sequence
c	356.2	30.1	66741	9	HS460J8	AL031662 Human DNA
19	356.2	30.1	145833	2	AC026539	AC026539 Homo sapi
20	212.6	18.0	1571	5	BC056035	BC056035 Xenopus l
21	159	13.4	2021	9	HSU44403	U44403 Human Src-1
22	158.6	13.4	3090	9	HSU44403	HSU44403 Homo sapi
23	157.4	13.3	1076	9	HSU30473	HSU30473 Sequence
24	157.4	13.3	1076	9	BC007042	BC007042 Homo sapi
25	157.4	13.3	1076	9	BC007042	BC007042 Homo sapi
26	157.4	13.3	2665	6	AX333017	AX333017 Sequence
27	157.4	13.3	2665	9	D89077	D89077 Homo sapien
c	156.2	13.2	145068	9	HSU977B1	AL050318 Human DNA
29	156.2	13.2	145833	2	AC026539	AC026539 Homo sapi
c	137.6	11.6	849	5	AY278230	AY278230 Gallus Ga
31	132	11.2	841	10	AY079449	AY079449 Mus muscu
32	132	11.2	841	10	AY079450	AY079450 Mus muscu
33	132	11.2	921	10	MMU29056	MMU29056 Mus muscu
34	132	11.2	2648	10	BCU32922	BCU32922 Mus muscu
35	132	11.2	2773	10	MMU131777	MMU131777 Mus muscu
c	129.8	11.0	299947	2	AC125701	AC125701 Rattus no
37	129.8	11.0	322940	2	AC123560	AC123560 Rattus no
38	128.8	10.9	1631	10	AY217759	AY217759 Rattus no
c	128.8	10.9	163240	10	AL935150	AL935150 Mouse DNA
40	126	10.7	3701	5	CHTKCKL	J03579 chicken tyr
41	126	10.7	4073	5	GGTKI	X60380 G.gallus tk
42	125.2	10.6	2703	10	MUSLYNB	M57697 Mouse lyn B
43	125.2	10.6	2766	10	MUSLYNA	M57696 Mouse lyn A
44	123.6	10.4	2041	10	RNAF000300	AF000300 Rattus no
45	123.6	10.4	2044	10	MUSLYN	M64608 Mouse lyn p

## ALIGNMENTS

RESULT 1	AX443133	1183 bp	DNA	linear	PAT 02-JUL-2002
AX443133	Sequence 74 from Patent WO0216599.				
LOCUS	AX443133				
DEFINITION	Sequence 74 from Patent WO0216599.				
ACCESSION	AX443133				
VERSION	AX443133.1	GI:21690555			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE					
AUTHORS	Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.				

TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES Location/Qualifiers  
source 1..1183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1183; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 1.2e-305;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTAGAGCTCAAGGACCCAGGCTGTGTCTCTGTGACAGAGCTCAAGAGGCGCTGGG 60  
Db 1 AGCTAGAGCTCAAGGACCCAGGCTGTGTCTCTGTGACAGAGCTCAAGAGGCGCTGGG 60

QY 61 CTTTCCCTCCCTGGCTGGCTGTGTCTGGAGGGTTCCCAAGTCCAGATCCCTAAGGAG 120  
Db 61 CTTTCCCTCCCTGGCTGGCTGTGTCTGGAGGGTTCCCAAGTCCAGATCCCTAAGGAG 120

QY 121 CATGGGCGAGCTGATCCATCCCTGTGTGACAACTGTGATGTCAGACAGATGCTGAGCT 180  
Db 121 CATGGGCGAGCTGATCCATCCCTGTGTGACAACTGTGATGTCAGACAGATGCTGAGCT 180

QY 181 ACCCAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAGTTCTGGGTG 240  
Db 181 ACCCAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAGTTCTGGGTG 240

QY 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAAGCTGTCCA 300  
Db 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAAGCTGTCCA 300

QY 301 GCAGAGCATGGTCTCAGCAGAGCTGCTCCCAAGCTTTGATGACAAACCAATTTC 360  
Db 301 GCAGAGCATGGTCTCAGCAGAGCTGCTCCCAAGCTTTGATGACAAACCAATTTC 360

QY 361 CTCGATGATGCTCTCTGAGTCTCTGCTGAGGAACTGGAAGTCTGCCAGCAGAG 420  
Db 361 CTCGATGATGCTCTCTGAGTCTCTGCTGAGGAACTGGAAGTCTGCCAGCAGAG 420

QY 421 AAAATCTCTGCAAGCCAGCTTGAGTCTCTCTGCAAGCCAGGACCTGTGACCAT 480  
Db 421 AAAATCTCTGCAAGCCAGCTTGAGTCTCTCTGCAAGCCAGGACCTGTGACCAT 480

QY 481 GGAAGCAGAGAGAAAGCAAGCCAGCTGGCCCTGGGAGTTTCCCGGAGGTGGCC 540  
Db 481 GGAAGCAGAGAGAAAGCAAGCCAGCTGGCCCTGGGAGTTTCCCGGAGGTGGCC 540

QY 541 GGCAGAGCTGTGCTGAGACTCGGGAGCCATGACCATCTGCTCTGAGGATGGAGACTG 600  
Db 541 GGCAGAGCTGTGCTGAGACTCGGGAGCCATGACCATCTGCTCTGAGGATGGAGACTG 600

QY 601 GTGGACGCTGTCTGTAAGTCTCAGGACAGAGTATTAACATCCCAAGCTCCAGTGGG 660  
Db 601 GTGGACGCTGTCTGTAAGTCTCAGGACAGAGTATTAACATCCCAAGCTCCAGTGGG 660

QY 661 CAAAGTCTCCATGGGTGCTGTATGAGGCTGTGACAGGGGAAAGCAGAGGAAGTCTCT 720  
Db 661 CAAAGTCTCCATGGGTGCTGTATGAGGCTGTGACAGGGGAAAGCAGAGGAAGTCTCT 720

QY 721 GTTGTACTCTGGAAACCTTGGAGGGCTTCTCTCATCCGGAGAGCCAGCAGAGAGG 780  
Db 721 GTTGTACTCTGGAAACCTTGGAGGGCTTCTCTCATCCGGAGAGCCAGCAGAGAGG 780

QY 781 CTCCTTACTCTGTAGTCCGCTCAGCCGCTGATCCTGGGACCGGATCAGACACTA 840  
Db 781 CTCCTTACTCTGTAGTCCGCTCAGCCGCTGATCCTGGGACCGGATCAGACACTA 840

QY 841 CAGGATCCACTGCTTGAATGGCTGGCTGTACATCTCACCGCGCTTCACTTCCCTC 900  
Db 841 CAGGATCCACTGCTTGAATGGCTGGCTGTACATCTCACCGCGCTTCACTTCCCTC 900

QY 901 ACTCCAGGCGCTGGTGAACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCAA 960  
Db 901 ACTCCAGGCGCTGGTGAACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCAA 960

QY 961 GGAGCCCTGTGTCTCTGACAGAGGCTGGCCGCTCCCTGCAAGGATATACCCCTACTCT 1020  
Db 961 GGAGCCCTGTGTCTCTGACAGAGGCTGGCCGCTCCCTGCAAGGATATACCCCTACTCT 1020

QY 1021 GACTGTGACAGAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTTCTGA 1080  
Db 1021 GACTGTGACAGAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTTCTGA 1080

QY 1081 AGCTGACACAGGGGAGAGTCTCTTCTCASTGAGGCTTCCGGAGTCCCTGAGTCTTA 1140  
Db 1081 AGCTGACACAGGGGAGAGTCTCTTCTCASTGAGGCTTCCGGAGTCCCTGAGTCTTA 1140

QY 1141 CATCAGCCTGAATGACAGAGGCTGTCTCTTTGGATGATGCTAG 1183  
Db 1141 CATCAGCCTGAATGACAGAGGCTGTCTCTTTGGATGATGCTAG 1183

RESULT 2  
AX443135/c  
LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002  
DEFINITION Sequence 76 from Patent WO0216599.  
ACCESSION AX443135  
VERSION AX443135.1 GI:21690556  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Burgess, C.E., Conley, P.B., Grosse, W.W., Hart, M., Kekuda, R., Shimkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E., Topper, J.N. and Yang, R.B.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES Location/Qualifiers  
source 1..1183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1183; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 1.2e-305;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTAGAGCTCAAGGACCCAGGCTGTGTCTCTGTGACAGAGCTCAAGAGGCGCTGGG 60  
Db 1183 AGCTAGAGCTCAAGGACCCAGGCTGTGTCTCTGTGACAGAGCTCAAGAGGCGCTGGG 1124

QY 61 CTTTCCCTCCCTGGCTGGCTGTGTCTGGAGGGTTCCCAAGTCCAGATCCCTAAGGAG 120  
Db 1123 CTTTCCCTCCCTGGCTGGCTGTGTCTGGAGGGTTCCCAAGTCCAGATCCCTAAGGAG 1064

QY 121 CATGGGCGAGCTGATCCATCCCTGGTGTACAACTGCTGCTGACAGACAGATGCTGAGCT 180  
Db 1063 CATGGGCGAGCTGATCCATCCCTGGTGTACAACTGCTGCTGACAGACAGATGCTGAGCT 1004

QY 181 ACCCAACCAACACCTAGCTCTCCCTGAGATCTCCAGGCTGAGAGAGTCTGGGTG 240  
Db 1003 ACCCAACCAACACCTAGCTCTCCCTGAGATCTCCAGGCTGAGAGAGTCTGGGTG 944

QY 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAAGCTGTCCA 300  
Db 943 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAAGCTGTCCA 884

QY 301 GCAGAGCATGCTCTCAGCAGAGCTGTCTTCCAGGCTTTCATGACAAACCAATTTC 360  
Db 301 GCAGAGCATGCTCTCAGCAGAGCTGTCTTCCAGGCTTTCATGACAAACCAATTTC 360



Db	858	CAGATCCACTGCTTGGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTC	917
Qy	901	ACTCCAGGCTGGTGGACCACTACTGAGCTGGGGATGACATCTGCTGCTACTCAA	960
Db	918	ACTCCAGGCTGGTGGACCACTACTGAGCTGGGGATGACATCTGCTGCTACTCAA	977
Qy	961	GGAGCCCTGTGCTCTGTCAGAGGGCTGGCCGCTCCCTGGCAAGATATACCCCTACTGT	1020
Db	978	GGAGCCCTGTGCTCTGTCAGAGGGCTGGCCGCTCCCTGGCAAGATATACCCCTACTGT	1037
Qy	1021	GACTGTGTCAGAGGACACCACTCACTGGAAGAGCTGGAGCTCCCTCTGTTTCTGA	1080
Db	1038	GACTGTGTCAGAGGACACCACTCACTGGAAGAGCTGGAGCTCCCTCTGTTTCTGA	1097
Qy	1081	AGCTGCCACAGGGAGGAGTCTCTTCTCAGTGGGGTCTCCGGAGTCCCTCAGCTCTTA	1140
Db	1098	AGCTGCCACAGGGAGGAGTCTCTTCTCAGTGGGGTCTCCGGAGTCCCTCAGCTCTTA	1157
Qy	1141	CATCAGCTGAATCAGAGGCTGTCTCTTCTGAGTGGATGCTAG	1193
Db	1158	CATCAGCTGAATCAGAGGCTGTCTCTTCTGAGTGGATGCTAG	1200
RESULT 4			
LOCUS	AX780857	2788 bp	DNA linear PAT 14-JUL-2003
DEFINITION	Sequence 3014 from Patent WO03039443.		
ACCESSION	AX780857		
VERSION	AX780857.1 GI:32697851		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	1		
AUTHORS	Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,		
TITLE	Dugas, M., Ellis, R., Brots, B. and Mergenthaler, S.		
JOURNAL	Novel genetic markers for leukemias		
FEATURES	Parent: WO 03039443-A 3014 15-MAY-2003;		
source	Deutsches Krebsforschungszentrum (DE) ;		
ORIGIN	Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,		
Query Match	PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)		
Best Local Similarity	1. .2788		
Matches 1153; Conservative	/organism="Homo sapiens"		
0; Mismatches 19; Indels 0; Gaps 0;	/mol_type="unassigned DNA"		
0;	/db_xref="taxon:9606"		
Qy	12	CAAGGACCCAGCGCTGTCTCTGTGACAGAGCTCAAGGGCCCTGGGCTTCCCTCCC	71
Db	1	CAAGGACCCAGCGCTGTCTCTGTGACAGAGCTCAAGGGCCCTGGGCTTCCCTCCC	60
Qy	72	TGGCTCGGCTGTCTTGGAGGGTTCCCGAGTCCAGATCCCTAAGAGCATGGGGCAGC	131
Db	61	TGGCTCGGCTGTCTTGGAGGGTTCCCGAGTCCAGATCCCTAAGAGCATGGGGCAGC	120
Qy	132	TGATCATCCCTGGTGTACAACTGCTGCTGACACAGATGCTGAGTACCACCAACAA	191
Db	121	TGATCATCCCTGGTGTACAACTGCTGCTGACACAGATGCTGAGTACCACCAACAA	180
Qy	192	CACCTAGCTCTCCCTGGAAGATCCCTCCAGGCTGAGAGATTTCTGGGTGTCTAGGACA	251
Db	181	CACCTAGCTCTCCCTGGAAGATCCCTCCAGGCTGAGAGATTTCTGGGTGTCTAGGACA	240
Qy	252	AGGACTGGCAGACTTCCAGAGGGCCCGCAAGCCCTAACTCTCCAGCCAGAGCATG	311
Db	241	AGGACTGGCAGACTTCCAGAGGGCCCGCAAGCCCTAACTCTCCAGCCAGAGCATG	300

Qy	312	CGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCGATGATG	371
Db	301	CGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCGATGATG	360
Qy	372	GCTTCTGAGTCTCTGCTGAGGAACAATGGGAAGTCTGCCAGCAGAGAAGAAATCTCTGC	431
Db	361	GCTTCTGAGTCTCTGCTGAGGAACAATGGGAAGTCTGCCAGCAGAGAAGAAATCTCTGC	420
Qy	432	CAAGCCCAAGCTTGAGTTCTCTGTCAGGCCAGGACCTGTACCATGGAAGCAGAGA	491
Db	421	CAAGCCCAAGCTTGAGTTCTCTGTCAGGCCAGGACCTGTACCATGGAAGCAGAGA	480
Qy	492	GAAGCAAGCCACAGCCGCTGGGCACTTTCCGGCAGAGTGGCCCGCCCGAGCTGT	551
Db	481	GAAGCAAGCCACAGCCGCTGGGCACTTTCCGGCAGAGTGGCCCGCCCGAGCTGT	540
Qy	552	CGCTCAGACTCGGGAGCCATGACCATCTCTCTGAGGATGGAGACTGTGTGGAGGTC	611
Db	541	CGCTCAGACTCGGGAGCCATGACCATCTCTCTGAGGATGGAGACTGTGTGGAGGTC	600
Qy	612	TGCTTGAAGTCTCAGGCAGAGAGTATAACATCCCGCAGCCTCCAGTGGGCAAGTCTCCC	671
Db	601	TGCTTGAAGTCTCAGGCAGAGAGTATAACATCCCGCAGCCTCCAGTGGGCAAGTCTCCC	660
Qy	672	ATGGGTGGCTGTATGAGGCTGTGACAGGGAGAAAGCAGAGAACTGTGTGTACTCTG	731
Db	661	ATGGGTGGCTGTATGAGGCTGTGACAGGGAGAAAGCAGAGAACTGTGTGTACTCTG	720
Qy	732	GAAGCCCTGGAGGGCTTCTCTATCCCGGAGAGCCAGACACAGAGAGGCTTCTTACTCTC	791
Db	721	GAAGCCCTGGAGGGCTTCTCTATCCCGGAGAGCCAGACACAGAGAGGCTTCTTACTCTC	780
Qy	792	TGTCAGTCCGCTCAGCCGCTGTCATCTCTGGAGCCGATCAGACACTACAGGATCCACT	851
Db	781	TGTCAGTCCGCTCAGCCGCTGTCATCTCTGGAGCCGATCAGACACTACAGGATCCACT	840
Qy	852	GCTTGCAGATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCACCTCCAGGCCC	911
Db	841	GCTTGCAGATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCACCTCCAGGCCC	900
Qy	912	TGCTGACCATTAATCTGAGCTGGCGGATGACATCTGTGTGCTACTCAAGAGAGCCCTGTG	971
Db	901	TGCTGACCATTAATCTGAGCTGGCGGATGACATCTGTGTGCTACTCAAGAGAGCCCTGTG	960
Qy	972	TCTTCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGA	1031
Db	961	TCTTCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGA	1020
Qy	1032	GGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTTCTTGAAGCTGGCACAG	1091
Db	1021	GGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTTCTTGAAGCTGGCACAG	1080
Qy	1092	GGAGAGGCTCTTCTCAGTGGGCTTCCGGAGTCCCTCAGCTTCTACATCAGCCTGA	1151
Db	1081	GGAGAGGCTCTTCTCAGTGGGCTTCCGGAGTCCCTCAGCTTCTACATCAGCCTGA	1140
Qy	1152	ATCAGCAGGCTGTCTCTTGGATGATGCCTAG	1183
Db	1141	ATCAGCAGGCTGTCTCTTGGATGATGCCTAG	1172

RESULT 5  
BC042041 2538 bp mRNA linear PRI 07-OCT-2003  
LOCUS  
DEFINITION  
Homo sapiens Szc-like-adaptor 2, transcript variant 1, mRNA (cdna  
clone MSC:49845 IMAGE:4429896), complete cds.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BC042041.1 GI:27469842  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;





```
781 CATCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAACATGGCTGCTGTACA 840
786 TCTCACCGCCCTCACCTCCCTCACTCCAGGCCCTGGTGACCACTTACTCTGAGCTGG 935
841 TCTCACCGCCCTCACCTCCCTCACTCCAGGCCCTGGTGACCACTTACTCTGAGCTGG 900
936 CGGATGACATCTGCTGCTACTCAAGAGAGCCCTGTGTCTTCGAGAGGCTGGCCGCTCC 995
901 CGGATGACATCTGCTGCTACTCAAGAGAGCCCTGTGTCTTCGAGAGGCTGGCCGCTCC 960
996 CTGCAAGGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCAACTGGAAGAGC 1055
961 CTGCAAGGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCAACTGGAAGAGC 1020
1056 TGGACAGCTCCCTCTGCTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGG 1115
1021 TGGACAGCTCCCTCTGCTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGG 1080
1116 GTCTCCGGGAGTCCCTGAGCTTCTACATCAGCTGGAATGAGGAGCTGTCTCTTGGATG 1175
1081 GTCTCCGGGAGTCCCTGAGCTTCTACATCAGCTGGAATGAGGAGCTGTCTCTTGGATG 1140
1176 ATGCTCTAG 1183
1141 ATGCTCTAG 1148

RESULT 6
AK025645 2415 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA: FLJ21992 fis, clone HEP06554.
DEFINITION AK025645
ACCESSION AK025645
VERSION AK025645.1 GI:10438227
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y.,
Okamoto S., Okitani R., Oai T., Suzuki Y., Obayashi M., Nishi T.,
Shibahara T., Tanaka T., Nakamura Y., Isogai T. and Sugano S.
TITLE NEPO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2415)
Sugano S., Suzuki Y., Oai T., Obayashi M., Nishi T., Isogai T.,
Shibahara T., Tanaka T. and Nakamura Y.
Direct Submission
TITLE Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)
COMMENT NEPO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- and 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES Location/Qualifiers
source 1..2415
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
clone="HEP06554"
cell_line="HepG2"
cell_type="hepatoma"
clone_lib="HEP"
note="cloning vector pME18SFL3"
1..846
note="unnamed protein product"
```

CDS

```
/codon_start=1
/protein_id="BAB15201.1"
/db_xref="GI:10438228"
/translation="MGSLLPSSRRKSLPSPLSSSSVQGGQPVMTAEKSKATAVALGSPF
AGGPAELSLRLGPEPLTIVSDGDWMTVLSEVSGREYNI PSVHAKVSHGWLVEGLSSE
KAEELLNPNPGGAFILRSOTERGSYLSVRLSRPASWDRI RHVRIHCLDNGWLVI
SPRLTFPSLQALVDHYSELADIDICLLKPCVLQKAGPLPKGDIPLPVTQRTLNWK
ELDSSLFSEAAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"

ORIGIN
Query Match 69.8%; Score 825.8; DB 9; Length 2415;
Best Local Similarity 99.8%; Pred. No. 5.4e-210;
Matches 827; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 355 ATTTCCTCCGATGATGCTCTTCGAGTCTCTCTGAGGAACAATGGGAAGTCTGCCAG 414
Db 18 ATTTCCTCCGATGATGCTCTTCGAGTCTCTCTGAGGAACAATGGGAAGTCTGCCAG 77
QY 415 CAGAAAGAAATCTCTGCCAAGCCCAAGCTTCAGATTCTCTCTCCAAAGCCAGGACCTGT 474
Db 78 CAGAAAGAAATCTCTGCCAAGCCCAAGCTTCAGATTCTCTCTCCAAAGCCAGGACCTGT 137
QY 475 GACCATGGAAGCAGAGAGAGCAAGCCACAGCCGTGCGCTGGCAGTTTCCCGGAGG 534
Db 138 GACCATGGAAGCAGAGAGAGCAAGCCACAGCCGTGCGCTGGCAGTTTCCCGGAGG 197
QY 535 TGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCTCTAGGATGG 594
Db 198 TGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCTCTAGGATGG 257
QY 595 AGACTGTGGAGCGGTGCTGTGAAAGTCTCAGGCGAGAGATATAACATCCCGAGGCTCA 654
Db 258 AGACTGTGGAGCGGTGCTGTGAAAGTCTCAGGCGAGAGATATAACATCCCGAGGCTCA 317
QY 655 CGTGGGCAAGCTCCCATGGTGGCTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGA 714
Db 318 CGTGGCAGAAAGTCTCCATGGTGGCTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGA 377
QY 715 ACTGCTGTGTTACCTGGGAAACCCCTGGAGGGCCCTTCCTCATCTCCGGAGAGCCAGACC 774
Db 378 ACTGCTGTGTTACCTGGGAAACCCCTGGAGGGCCCTTCCTCATCTCCGGAGAGCCAGACC 437
QY 775 GAGAGCTCTTACTCTCTGAGTCCGCTCAGCCGCTGATCTCTGGACCGGATCAG 834
Db 438 GAGAGCTCTTACTCTCTGAGTCCGCTCAGCCGCTGATCTCTGGACCGGATCAG 497
QY 835 AACTACAGGATCCACTGCTTTGACATGGCTGGCTGTACATCTCACCGGCTCACCTT 894
Db 498 AACTACAGGATCCACTGCTTTGACATGGCTGGCTGTACATCTCACCGGCTCACCTT 557
QY 895 CCCTCACTCCAGGCCCTGGTGAGCAATTAATCTGAGCTGGCGGATGACATCTGCTGCT 954
Db 558 CCCTCACTCCAGGCCCTGGTGAGCAATTAATCTGAGCTGGCGGATGACATCTGCTGCT 617
QY 955 ACTCAAGAGAGCCCTGTGCTCTGCAGAGGCGCTGGCCCGCTCCCTGGCAAGGATATACCC 1014
Db 618 ACTCAAGAGAGCCCTGTGCTCTGCAGAGGCGCTGGCCCGCTCCCTGGCAAGGATATACCC 677
QY 1015 ACCTGTGACTGTGCAGAGGACACCACTCACTGGAAGAGCTGGAGAGCTCCCTCTGTT 1074
Db 678 ACCTGTGACTGTGCAGAGGACACCACTCACTGGAAGAGCTGGAGAGCTCCCTCTGTT 737
QY 1075 TTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAG 1134
Db 738 TTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAG 797
QY 1135 CTCTCATCATCAGCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183
Db 798 CTCTCATCATCAGCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 846

RESULT 7
AXS11153
```

LOCUS AX511153 786 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 4 from Patent WO0242452.  
ACCESSION AX511153  
VERSION AX511153.1 GI:23392046  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Mcglade,J.C. and Loreto,M.P.  
TITLE Adapter gene  
JOURNAL Patent: WO 0242452-A 4 30-MAY-2002;  
The Hospital for Sick Children (CA)  
FEATURES  
source  
1..786  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 66.3%; Score 784.4; DB 6; Length 786;  
Best Local Similarity 99.9%; Pred. No. 6.9e-199;  
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTGAGTTCCTCTGTC 457  
DB 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTGAGTTCCTCTGTC 60  
QY 458 CAAGCCAGGACCTGTGACCATGAGCAGAGAGAAAGCCACAGCGTGGCCCTG 517  
DB 61 CAAGCCAGGACCTGTGACCATGAGCAGAGAGAAAGCCACAGCGTGGCCCTG 120  
QY 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGCGTGAGACTCGGGAGCCATTGACC 577  
DB 121 GCGAGTTTCCCGCAGGTGGCCCGCCGAGCTGCGTGAGACTCGGGAGCCATTGACC 180  
QY 578 ATCGTCTCTGAGGATGAGACTGTGGACGGTGTGCTGTGAAGTCTCAGGCAGAGATAT 637  
DB 181 ATCGTCTCTGAGGATGAGACTGTGGACGGTGTGCTGTGAAGTCTCAGGCAGAGATAT 240  
QY 638 AACATCCCAGGTCACGTCGGCAAGTCTCCCATGGTGGTGTATGAGGGCTGAGC 697  
DB 241 AACATCCCAGGTCACGTCGGCAAGTCTCCCATGGTGGTGTATGAGGGCTGAGC 300  
QY 698 AGGAGAAAGCAGAGAACTGTGTGTTTACTTGGGAACCTTGGAGGGCTTCTCTCATC 757  
DB 301 AGGAGAAAGCAGAGAACTGTGTGTTTACTTGGGAACCTTGGAGGGCTTCTCTCATC 360  
QY 758 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTGTCAGTCGCGCTCAGCGCCCTGCA 817  
DB 361 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTGTCAGTCGCGCTCAGCGCCCTGCA 420  
QY 818 TCCTGGGACCGGATCAGACATACAGGATCCACTTGCCTTGACATGGCTGCTGTACATC 877  
DB 421 TCCTGGGACCGGATCAGACATACAGGATCCACTTGCCTTGACATGGCTGCTGTACATC 480  
QY 878 TCACCGCCCTCAGCTTCCCTCTACTCCAGCCCTGTGGACCAATTAATCTGAGCTGGCG 937  
DB 481 TCACCGCCCTCAGCTTCCCTCTACTCCAGCCCTGTGGACCAATTAATCTGAGCTGGCG 540  
QY 938 GATGACATCTGCTGCTACTCAGAGAGCCCTGTGCTGTCAGAGGGCTGGCCGCTCCCT 997  
DB 541 GATGACATCTGCTGCTACTCAGAGAGCCCTGTGCTGTCAGAGGGCTGGCCGCTCCCT 600  
QY 998 GCGAAGGATATACCCCTACCTCTGTCAGAGGACACCACTCAACTGGAAGAGCTG 1057  
DB 601 GCGAAGGATATACCCCTACCTCTGTCAGAGGACACCACTCAACTGGAAGAGCTG 660  
QY 1058 GACAGCTCCCTCTGTTTCTGGAAGCTGCCAGGGAGGAGTCTCTCTCAGTGAGGGT 1117  
DB 661 GACAGCTCCCTCTGTTTCTGGAAGCTGCCAGGGAGGAGTCTCTCTCAGTGAGGGT 720

QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGAGAGGCTGTCTCTTTGATGAT 1177  
DB 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGAGAGGCTGTCTCTTTGATGAT 780  
QY 1178 GCCTAG 1183  
DB 781 GCCTAG 786  
RESULT 8  
LOCUS AX572845 786 bp DNA linear PAT 29-NOV-2002  
DEFINITION Sequence 1 from Patent WO02055707.  
ACCESSION AX572845  
VERSION AX572845.1 GI:26004935  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Holland,S.J., Mendenhall,M.K., Pardo,J., Spencer,C., Fu,A.C.,  
Luo,Y., Payan,D.G., Mancebo,H.S., Wu,J., Zhou,X., Shen,M.,  
Liao,X.C. and Sheng,N.  
TITLE Cloning of an inhibitor of antigen-receptor signaling by a  
retroviral-based functional screen  
JOURNAL Patent: WO 02055707-A 1 18-JUL-2002;  
Rigel Pharmaceuticals, Inc. (US)  
FEATURES  
source  
1..786  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..786  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD58542.1"  
/db\_xref="GI:26004936"  
/db\_xref="REMBL:CAD58542"  
/translation="MGSLPSRKSLSPSLSSVGGPGVTMEERSKATAVALGSPF  
AGGPELSRLGEPLTVSEDDGAWTVLSEVSGREYNIPSHVAKVSHGWLVEGLSRE  
KABELLLPQNPQGAFLIRSEQTRGYSLSVRLSRPASWDRIHYRHLCDNGWLYI  
SPRLTFQALVDHYLSLADDDICCLKEPCVQLRAGPLFGKDIPLPVTQVRIPLNWK  
ELDSSLPLSEATGEBESLLSEGLRESLSFYISLNDEAVSLDDA"  
ORIGIN  
Query Match 66.3%; Score 784.4; DB 6; Length 786;  
Best Local Similarity 99.9%; Pred. No. 6.9e-199;  
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTGAGTTCCTCTGTC 457  
DB 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTGAGTTCCTCTGTC 60  
QY 458 CAAGCCAGGACCTGTGACCATGAGCAGAGAGAAAGCCACAGCGTGGCCCTG 517  
DB 61 CAAGCCAGGACCTGTGACCATGAGCAGAGAGAAAGCCACAGCGTGGCCCTG 120  
QY 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGCGTGAGACTCGGGAGCCATTGACC 577  
DB 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGCGTGAGACTCGGGAGCCATTGACC 180  
QY 578 ATCGTCTCTGAGATGAGACTGTGGACGGTGTGCTGTGAAGTCTCAGGCAGAGATAT 637  
DB 181 ATCGTCTCTGAGATGAGACTGTGGACGGTGTGCTGTGAAGTCTCAGGCAGAGATAT 240  
QY 638 AACATCCCAGGTCACGTCGGCAAGTCTCCCATGGTGGTGTATGAGGGCTGAGC 697  
DB 241 AACATCCCAGGTCACGTCGGCAAGTCTCCCATGGTGGTGTATGAGGGCTGAGC 300  
QY 698 AGGAGAAAGCAGAGAACTGTGTGTTTACTTGGGAACCTTGGAGGGCTTCTCTCATC 757  
DB 301 AGGAGAAAGCAGAGAACTGTGTGTTTACTTGGGAACCTTGGAGGGCTTCTCTCATC 360



[illegible]

Db 241 AACATCCCAAGCTCCACGTCGGCAAGTCTCCATGGGTGGTGTATGAGGCGCTGAGC 300

Qy 698 AGGAGAGAAACAGAGAACTGCTGTGTTTACCTGGGAACCTCGAGGGGCTTCTCTCATC 757

Db 301 AGGAGAGAAACAGAGAACTGCTGTGTTTACCTGGGAACCTCGAGGGGCTTCTCTCATC 360

Qy 758 GGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTCTCAGTCGCTCAGCCGCTGCA 817

Db 361 CGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTCTCAGTCGCTCAGCCGCTGCA 420

Qy 818 TCCTGGACCGGATCAGACACTCAGGATCCACTGCTTGCATCAATGGCTGTACATC 877

Db 421 TCCTGGACCGGATCAGACACTCAGGATCCACTGCTTGCATCAATGGCTGTACATC 480

Qy 878 TCACCGGCTCAGCTCCCTCCTACTCCAGGCGCTGTGAGCACTTACTCTGAGCTGGG 937

Db 481 TCACCGGCTCAGCTCCCTCCTACTCCAGGCGCTGTGAGCACTTACTCT 531

Qy 938 GATGACATCTGCTCCTACTCAAGAGCCCTGTGTCTCTGACAGGGGCTGGCCGCTCCCT 997

Db 532 -----GAGGGCTGGCCGCTCCCT 550

Qy 998 GCGAAGATATACCTTACCTGTGACTGTGAGAGGACACACTCACTCGAAGAGCTG 1057

Db 551 GCGAAGATATACCTTACCTGTGACTGTGAGAGGACACACTCACTCACTGGAAGAGCTG 610

Qy 1058 GACAGCTCCCTCCCTGTTTCTGAGCTGCCACAGGGGAGGAGTCTTCTCAGTGAAGGT 1117

Db 611 GACAGCTCCCTCCCTGTTTCTGAGCTGCCACAGGGGAGGAGTCTTCTCAGTGAAGGT 670

Qy 1118 CTCGGAGTCCCTCAGCTTCTATCAGCTGAAATGA-CGAGGCTGTCTTTGGATGA 1176

Db 671 CTCGGAGTCCCTCAGCTTCTATCAGCTGAAATGA-CGAGGCTGTCTTTGGATGA 730

Qy 1177 TGCCTAG 1183

Db 731 TGCCTAG 737

RESULT 12

LOCUS AF290986 737 bp mRNA linear PRI 21-JAN-2003

DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA, complete cds; alternatively spliced.

ACCESSION AF290986

VERSION AF290986.1 GI:17351922

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 737)

AUTHORS Loreto,M.P. and McGlade,C.J.

TITLE Cloning and characterization of human Src-like adaptor protein 2 and a novel splice isoform, SLAP-2-v

JOURNAL Oncogene 22 (2), 266-273 (2003)

MEDLINE 22415750

PUBMED 12527895

REFERENCE 2 (bases 1 to 737)

AUTHORS Loreto,M.P. and McGlade,C.J.

TITLE Direct Submission

JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada

FEATURES

source location/Qualifiers

1..737

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="20"

/tissue\_type="thymus"

1..633

/note="SLAP-2-v"

/codon\_start=1

CDS

/product="Src-like adaptor protein-2 splice isoform"

/protein\_id="AAL38198.1"

/db\_xref="GI:17351923"

/translation="MGSLLPSSRRKSLPSPSSVQGGQPPVMEAKSKATAVALGSPF AGPAELSLRLGPLETIVSDGWMTVLSEVSGREYNIPSHVAKSHGWLKYLSE KAEELLPLGNPGGAFILRESQTRGYSLSVRLSPASWDRIHRIHCLDNGWLYI SPRLTFPSLQALVDHYSEGHFAPWQGYTPTCDAEDTTQLERAGLPPEV"

ORIGIN

Query Match 55.9%; Score 661.8; DB 9; Length 737;

Best Local Similarity 93.3%; Pred. No. 4.7e-166;

Matches 734; Conservative 0; Mismatches 2; Indels 51; Gaps 2;

Qy 398 ATGGGAATCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 457

Db 1 ATGGGAATCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 60

Qy 458 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCGTGGCCCTG 517

Db 61 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCGTGGCCCTG 120

Qy 518 GCGATTTCCCGCAGGTGGCCCGCGAGCTGCTGAGACTCGGAGGCCATTGACC 577

Db 121 GCGATTTCCCGCAGGTGGCCCGCGAGCTGCTGAGACTCGGAGGCCATTGACC 180

Qy 578 ATCGTCTCTGAGGATGGAGACTGGTGGACGCTGTCTGTAAGTCTCAGGAGAGATAT 637

Db 181 ATCGTCTCTGAGGATGGAGACTGGTGGACGCTGTCTGTAAGTCTCAGGAGAGATAT 240

Qy 638 AACATCCCGCAGGCTCCAGTGGGCAAGTCTCCCATGGGTGGTGTATGAGGCGCTGAGC 697

Db 241 AACATCCCGCAGGCTCCAGTGGGCAAGTCTCCCATGGGTGGTGTATGAGGCGCTGAGC 300

Qy 698 AGGAGAGAAAGCAGAGAGAACTGCTGTGTTTACCTGGGAACCTCGAGGGGCTTCTCTATC 757

Db 301 AGGAGAGAAAGCAGAGAGAACTGCTGTGTTTACCTGGGAACCTCGAGGGGCTTCTCTATC 360

Qy 758 CGGAGAGCCAGACAGGAGAGGCTTCTCTCTGTCAGTCCGCTCAGCGGCCCTGCA 817

Db 361 CGGAGAGCCAGACAGGAGAGGCTTCTCTCTGTCAGTCCGCTCAGCGGCCCTGCA 420

Qy 818 TCCTGGGACCGGATCAGACACTCAGGATCCACTGCTTGCATCAATGGCTGTACATC 877

Db 421 TCCTGGGACCGGATCAGACACTCAGGATCCACTGCTTGCATCAATGGCTGTACATC 480

Qy 878 TCACCGGCTCAGCTCCCTCCTACTCCAGGCGCTGTGAGCACTTACTCTGAGCTGGG 937

Db 481 TCACCGGCTCAGCTCCCTCCTACTCCAGGCGCTGTGAGCACTTACTCT 531

Qy 938 GATGACATCTGCTGCTACTCAAGAGCCCTGTGTCTCTGACAGGGGCTGGCCGCTCCCT 997

Db 532 -----GAGGGCTGGCCGCTCCCT 550

Qy 998 GCGAAGATATACCTTACCTGTGACTGTGAGAGGACACACTCACTCGAAGAGCTG 1057

Db 551 GCGAAGATATACCTTACCTGTGACTGTGAGAGGACACACTCACTCACTGGAAGAGCTG 610

Qy 1058 GACAGCTCCCTCCTGTTTCTGAGCTGCCACAGGGGAGGAGTCTTCTCTCAGTGAAGGT 1117

Db 611 GACAGCTCCCTCCTGTTTCTGAGCTGCCACAGGGGAGGAGTCTTCTCTCAGTGAAGGT 670

Qy 1118 CTCGGAGTCCCTCAGCTTCTATCAGCTGAAATGA-CGAGGCTGTCTTTGGATGA 1176

Db 671 CTCGGAGTCCCTCAGCTTCTATCAGCTGAAATGA-CGAGGCTGTCTTTGGATGA 730

Qy 1177 TGCCTAG 1183

Db 731 TGCCTAG 737

RESULT 13

LOCUS BC052655 1384 bp mRNA linear ROD 04-NOV-2003

DEFINITION Mus musculus Src-like-adaptor 2, mRNA (cDNA clone MGC:60811 IMAGE:30040401), complete cds.

## ACCESSION

BC052655

## VERSION

BC052655.1

## KEYWORDS

MGC.

## SOURCE

Mus musculus

## ORGANISM

Mus musculus

## REFERENCE

1 (bases 1 to 1384)

## AUTHORS

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P., Dietzenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L., Schreier,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Kuzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fanej,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butcherfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

## TITLE

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

## MEDLINE

22388257

## PUBMED

12477932

## REFERENCE

2 (bases 1 to 1384)

## AUTHORS

Strausberg,R.

## TITLE

Direct Submission

## JOURNAL

Submitted (16-MAY-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

## COMMENT

Contact: MGC help desk

Email: [cgabp-help@mail.nih.gov](mailto:cgabp-help@mail.nih.gov)

Tissue Procurement: Drs. Dennis Taub, Dan Longo (NIA, USA), Jonathan Keller (NCI, USA)

cDNA Library Preparation: Yulan Piao and Minoru Ko (National Institute on Aging, NIH: <http://lgsun.grc.nia.nih.gov/cDNA/>)

cDNA Library Arrayed by: The i.M.A.G.E. Consortium (ILLN)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)

Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantiripop,S., Thomas,P.J., Touchman,J.W., Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the i.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>

Series: IRAP Plate: 112 Row: e Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

Location/Qualifiers

1..1384

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6Ncr"

## FEATURES

source

/db\_xref="taxon:10090"  
/clone="MGC:60811 IMAGE:30040401"  
/tissue\_type="Hematopoietic Stem Cell,  
(Lin-/C-Kit+/Sca-1-), mouse, 10 wks"  
/clone\_lib="NIA Mouse Hematopoietic Stem Cell  
(Lin-/C-Kit+/Sca-1-) cDNA Library (long)"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1"  
1..1384  
/gene="Slap2"  
/note="synonyms: SLAP2, SLAP-2"  
/db\_xref="LocusID:77799"  
/db\_xref="MGI:1925049"  
330..1109  
/codon\_start=1  
/product="Slap2 protein"  
/protein\_id="AAH52655.1"  
/db\_xref="GI:30851668"  
/db\_xref="LocusID:77799"  
/translation="MGSLSRGKTSPPSSSGDPQEPVSMOPERRKVTAVAGSPPA  
GEARLSRLBEPPLTIISEDDQWTVQSEVSGREYHMSVYAKVAHGLIEGLSREK  
AEELLLPFGAFILRESQTRRCISLVRLSPASWDRHRRHRLQRLNGWLYIS  
PLTTPFSLHALVHYSELAAGICCPRLPCLQKGLPFGKDPPTVPTVPTSSSLNKK  
LDRSLFLFAPASGEASLLSEGLRESLSYSLSAEDP.DDA"  
432..596  
/note="SH3; Region: SH3 domain. SH3 (Src homology 3)  
domains are often indicative of a protein involved in  
signal transduction related to cytoskeletal organisation.  
First described in the Src cytoplasmic tyrosine kinase.  
The structure is a partly opened beta barrel"  
/db\_xref="CDD:pfam00018"  
603..899  
/note="SH2; Region: Src homology 2 domains"  
/db\_xref="CDD:cd00173"

## ORIGIN

Query Match 47.6%; Score 563; DB 10; Length 1384;  
Best Local Similarity 78.3%; Pred. No. 1.3e-139;  
Matches 715; Conservative 0; Mismatches 190; Indels 8; Gaps 3;  
QY 274 AGGCGCCCAAGACCTTACCTGTCCAGCCAGCATGCTCAGCAGAGCTGTCTTCC 333  
DB 208 AGGACCTGCAAGGCTTACCTGTGGTCACTG-TGCACATTGGCTACTACCTCAT 265  
QY 334 CAAGCCTTTGATGACAAACCAATTTCCCTCGATGATGCTTCTGAGTCTCTGTGAGG 393  
DB 266 CAAACGCTGTATGGCAACCTTTCCCTTCCAGGTTCACTGTGTTGTGAGGCTGTGCTG 325  
QY 394 AACATGGGAAGTCTCCCGCAGCAGAGAAATCTCTGCCAGCCCAAGCTTCACTTCCCTC 453  
DB 326 AGTGATGGGAAGTTTGTCTCAGCAGAGGGAACCT---CCAGCCCAAGCCCAAGCTTCC 382  
QY 454 TGTCCAAGCCAGGACCTGTGACCATGGAACAGAGAGAACAGGCCACAGCCGTGGC 513  
DB 383 TGTCCAGACACAGGAACCCGTGTCATGCAACAGAGAGCAGCAAGGTCACTGAGTGGC 442  
QY 514 CTGGGCACTTTCCCGCAGGTTGGCCCGCCAGCTGTCTGAGACTCGGGAGGCCATT 573  
DB 443 CTGGGCACTTTCCCGCAGGTTGAACAGCCCAAGCTTCTGAGACTCGGGAGGCCGT 502  
QY 574 GACCATCGTCTCTGAGGATGGAGACTGTGTGGACGGTGTCTGTAAGTCTCAGGACAGAG 633  
DB 503 GACCATCATCTCTGAGGATGGAGATGTGTGGACAGTCCAGTCCGAAGTCTCAGGACAGAG 562  
QY 634 GTATACATCCCGAGGCTCCAGTGGGCAAGTCTCCCATGAGGTGGTGTGTATGAGGGCT 593  
DB 563 GTACACATGCCAGTGTGTATGTGGCTAAAGTCGCCCAAGTTCAGTGGGGGCT 622  
QY 694 GAGCAGGAGAGAAAGCAGAGGAATCTGCTGTGTACTGGGAACCTTGGAGGGGCTTCTCT 753  
DB 623 GAGCCGGGAGAAAGCCGAGGAATCTACTCTGTACTTGGGAACCCCGAGGGGCTTCTCT 682  
QY 754 CATCCGGGAGAGCCAGACACAGGAGAGGCTCTTACTCTCTGTCTGAGTCCGCGCTCAGCGGCC 813



```
Db 683 CATCCGGGAGAGCCAGACACAGAGAGGCTGCTATTCCCTCTCCGTCGAGCTAGCCGCC 742
Qy 814 TGCATCTCTGGGACCGGATCAGACACTACAGGATCCACTGCTTTTGACAATGGCTGGTGTA 873
Db 743 TGCATCTTGGGACCGGATCAGACACTACAGGATACAGCGTCTTTGACAATGGCTGGTGTA 802
Qy 874 CATCTCACCGGCTCAGCTCCCTCCCTCAGTCCAGGCGCTGGTGAGCAATCTCTGAGCT 933
Db 803 CATCTCACCTCTGCTCAGCTCCCTCCCTCAGTCCAGGCGCTGGTGAGCAATCTCTGAGCT 862
Qy 934 GCGGATCAGATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 993
Db 863 AGCAGATGGCATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 922
Qy 994 CCCTGGCAAGATATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1053
Db 923 ACCTGGCAAGATATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 982
Qy 1054 GCTGGAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1110
Db 983 GCTGGAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1042
Qy 1111 TGAGGCTCTCGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1170
Db 1043 TGAGGCTCTCGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1102
Qy 1171 GGATGATGCTCTAG 1183
Db 1103 TGCTTAGCCCTGG 1115

RESULT 14
AX511150
LOCUS AX511150 1348 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1 from Patent WO0242452.
ACCESSION AX511150
VERSION AX511150.1 GI:23392044
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS McGlade, J.C. and Loreto, M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 1 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source
1. 1348
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 47.5%; Score 561.4; DB 6; Length 1348;
Best Local Similarity 78.2%; Pred. No. 3.6e-139;
Matches 714; Conservative 0; Mismatches 191; Indels 8; Gaps 3;

Qy 274 AGGCGCCCCAAAGCCCTTAACCTGCTCCAGCAGAGCATGCTCTCAGCAGAGCTGTCTTCC 333
Db 160 AGGACCTGCAAAAGGCGTACCTGCTGGGTCAGTG--TGACATTTGGCTGACTACCTTCAT 217
Qy 334 CAAGCCTTTGATGACAAACCAATTTCCCTCGATGATGCTTCTGAGTGCTCTGCTGAGG 393
Db 218 CAAGCTCTGATGGCAACCTTTCCCTTCCAGTTTCTGAGTGCTGCTGAGCTGTCTG 277
Qy 394 AACATGGGAAGTGTGCCAGCAGAGAAATCTCTCCAAAGCCCAAGCTTGAGTTCTCTC 453
Db 278 AGTGATGGGAAGTGTGCCAGCAGAGAAACCT---CCAGCCCCAGCCCCAGCTCTCTC 334
Qy 454 TGTCAGGCCAGGAGCTGTGACCATGGAAGCAGAGAGAGCAAGCCACAGCCGCTGGC 513
```

```
Db 335 TGGTCCAGCAGGAAACCCGTGTCTCCATCCAACAGAAAGACACAAGTTCACAGCTGTGGC 394
Qy 514 CTTGGGCAAGTTTCCCGCAGAGTGGCCCGCGAGCTGTCTGCTGAGACTCGCGGGAGCAATT 573
Db 395 CTTGGGCAAGTTTCCCGCAGAGTGAACAGCGCAGACTATCTCTGAGACTCGCGGGAGCGCT 454
Qy 574 GACCATCTCTCTGAGGATGAGAGCTGGTGACAGCGTGTCTGCTGAAGTCTCTCAGCAGAGA 633
Db 455 GACCATCTCTCTGAGGATGAGAGTGGTGACAGTCCAGTCGGAGTCTCTCAGCAGAGA 514
Qy 634 GTATAACATCCCGCAGCGTCCACGTGGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGGCT 693
Db 515 GTACCACATGCCCAGTGTGTATGTGGCTAAAGTCCGCCACCGGTGGTGTATGAGGGCT 574
Qy 694 GAGCAGGAGAGAAAGCAGAGGAAGTGTCTTCTTACCTGGGAAACCTGTGAGGGGCTTCTCT 753
Db 575 GAGCCGGAGAGAAAGCCGAGGAAGTACTCTCTGTACCTGGGAAACCCCGAGGGGCTTCTCT 634
Qy 754 CATCCGGAGAGCCAGACAGAGAGGCTCTTACTCTCTGTCTGTCAGTCCGCTCTCAGCGGCC 813
Db 635 CATCCGGAGAGCCAGACAGAGAGGCTGTATTTCCCTGCTCCGCTCCGACTCAGCGGCC 694
Qy 814 TGATCTCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCATGCTGCTGTA 873
Db 695 TGATCTCTGGGACCGGATCAGACACTACAGGATCAGGCTCTTGCATGCTGCTGTA 754
Qy 874 CATCTCACCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 933
Db 755 CATCTCACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 814
Qy 934 GCGGATGACATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 993
Db 815 AGCAGATGGCATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 874
Qy 994 CCTCGCAAGGATATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1053
Db 875 ACCTGGCAAGATATACACTCTCCACTGTGACTGTGCAACATCATCATCTCTCTCTCTCT 934
Qy 1054 GCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1110
Db 935 GCTGACCGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 994
Qy 1111 TGAGGCTCTCGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1170
Db 995 TGAGGCTCTCGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1054
Qy 1171 GGATGATGCTCTAG 1183
Db 1055 TGCTTAGCCCTGG 1067

RESULT 15
AF287467
LOCUS AF287467 1348 bp mRNA linear ROD 03-JUN-2002
DEFINITION Mus musculus Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF287467
VERSION AF287467.1 GI:17351918
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Loreto, M.P., Berry, D.M. and McGlade, C.J.
TITLE Functional cooperation between c-Cbl and Src-like adaptor protein 2 in the negative regulation of T-cell receptor signaling
JOURNAL Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
MEDLINE 22022020
PUBMED 12024036
REFERENCE 2 (bases 1 to 1348)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for
```



Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada

# FEATURES

## source

```

1..1348
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/dev_stage="day 15 embryo"
282..1061
/notes="SLAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAL38196.1"
/db_xref="GI:17351919"
/translation="MGSLSRSGKTSPPSSSPDQEPVSMOPERHKVTAVALGSFPA
GEOARLSRLGPLELIISEDDMTVQSEVGRVHPSVYAKVAGHMLVEGLSREK
ABELLLPNPGAGLIRSEQRCYSLSVLSPASWDRIHRVRIORLNGWLYIS
PLTFPSLHMLVEHSLGICCPRIEPCVLPKLPFGKDPFPVFPVTPVTSLSNWKX
LDRSLFLFAPASGEASLSLEGRSLSSYISLAEDPLDDA"
381..551
/misc_feature
/note="Region: SH3 domain"
552..824
/misc_feature
/note="Region: SH2 domain"
1301..1306
/polyA_signal
ORIGIN

```

```

Query Match      47.5%; Score 561.4; DB 10; Length 1348;
Best Local Similarity 78.2%; Pred. No. 3.6e-139;
Matches 714; Conservative 0; Mismatches 191; Indels 8; Gaps 3;

QY 274 AGGGCCCCCAAGCCCTACCTGCTCCAGCAGAGATCGGTCTCAGCAGAGCTGTCTCC 333
DB 160 AGACCTGCAAGAGCCCTGACCTGTGGGTGCTG--TGACATTTGGCTGACTACCTCAT 217

QY 334 CAAGCCCTTTGATGACAAACCAATTTCCCTCGATGATGCTTCTGAGTGTCTGTGTAGG 393
DB 218 CAACCTGTGATGGCAACCTTTCCCTTTCCAGGTTCACTGTGCTGTGTGAGGCTCTGTG 277

QY 394 AACATGGGAAGTCTGCCAGCAGAGAAATCTGCAAGCCCAAGCTTGAAGTCTCTC 453
DB 278 AGTGATGGGAAGTTTGTCCAGCAGAGGAAACCT---CCAGCCCCAGCCCAAGCTCTCTC 334

QY 454 TGTCGAAGCCCAAGGACCTGTGACCATGGAAGCAGAGAAACCAAGGCCACAGCCGTGGC 513
DB 335 TGCTCCAGACCAAGAACCCGTGTCCATGCAACAGAAAGACACAGGTACAGCTGTGGC 394

QY 514 CCTGGCAGTTTCCGGCAGGTGGCCCGCCAGCTGTGCTGAGACTCGGGAGCCATT 573
DB 395 CCTGGCAGTTTCCAGCAGGTGAACAGGCCAGACTATCTCTGAGACTCGGGAGCCGCT 454

QY 574 GACCATCTCTCTGAGGATGGAGACTGTGGACGCTGTCTGAACTCTCAGGCAGAGA 633
DB 455 GACCATCTCTCTGAGGATGGAGATTGGTGGACAGTCCAGTCGGAAGTCTCAGGCAGAGA 514

QY 634 GTATAACATCCCAAGCTCCAGTGGGCAAGTCTCCATGGGTGGCTGTATGAGGGCCT 693
DB 515 GTACCACATGCCCAAGTGTATGTGGCTAAAGTCGCCACCGGTGGCTGTACGAGGGCCT 574

QY 694 GAGCAGGAGAAAGCAGAGAACTGCTGTGTGTAGCTGGACCCCTGGAGGGCCCTTCT 753
DB 575 GAGCCGGGAGAAAGCCAGAACTACTCTGTGTAGCTGGAAACCCCGAGGGGGCTTCTCT 634

QY 754 CATCCGGGAGAGCCAGACCAAGAGAGGCTTTTACTCTCTGTGCTCCGCTCAGCCGCC 813
DB 635 CATCCGGGAGAGCCAGACCAAGAGAGGCTGTATTCCCTGTCCGCTCCGACTCAGCCGCC 694

QY 814 TGCATCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACATGCTGCTGTA 873
DB 695 TGCATCTGGGACCGGATCAGACACTACAGGATCAGGCTCTTGACAATGCTGCTGTA 754

QY 874 CATCTCAGCCGCTCACCCTTCCCTCACTCCAGGCCCTTGGTGGACCATTTACTCTGAGCT 933
DB 755 CATCTCAGCTCGCTCACCCTTCCCTCACTCCAGGCCCTTGGTGGAGCATTTACTCTGAGCT 814

```

Search completed: February 19, 2004, 23:23:13  
Job time : 5187.29 secs

```

QY 934 GCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTCAGAGGCTGGCCGCT 993
DB 915 AGCAGATGGCATCTGCTGTCCCTCAGGAGCCGCTGTCTCTCAGAAAGCTTGGCCACT 874

QY 994 CCCTGGCAAGGATATACCCCTTACTGTGACTGTGACAGGACACCACTCAACTGGAAGA 1053
DB 875 ACCTGGCAAGGATATACCCCTTACTGTGACTGTGACAGGACACCACTCAACTGGAAGA 934

QY 1054 GCTGACAGCTCCCTCTCTGTTTCTGAAG---CTGCCACAGGGAGGAGTCTCTTCTCAG 1110
DB 935 GCTGACAGGAGCTCTCTGTTTCTGGAAGCACCTGCCAGTGGGAGGACATCTCTGCTCAG 994

QY 1111 TGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCTTCTACATCAGCTTCTCTCTTT 1170
DB 995 TGAGGGCTCCGAGAGTCCCTCAGTTCCTTACATCAGCTTCTCTCTCTGAGATGA 1054

QY 1171 GGATGATGCTAG 1183
DB 1055 TGCTTAGCCCTGG 1067

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:15:48 ; Search time 1096 Seconds  
(without alignment)  
4585.415 Million cell updates/sec

Title: US-09-939-853A-74  
Perfect score: 1183  
Sequence: 1 agctagagctccaaggacc.....tctcttgatgacgctag 1183

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N:Geneseq\_25Jan04.\*  
1: geneseqn1980s.\*  
2: geneseqn1990s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002s.\*  
7: geneseqn2003as.\*  
8: geneseqn2003bs.\*  
9: geneseqn2003cs.\*  
10: geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1183	100.0	1183	6	ABK61465 Human cDN
2	1181.4	99.9	2567	6	Ad43980 Human Src
3	784.4	66.3	786	6	Al440089 Human mod
4	784.4	66.3	786	6	Abq74343 Human Src
5	758.2	64.1	763	6	Abq98670 Human ORF
6	740.2	62.6	837	3	Abq77202 Human ORF
7	724.2	61.2	1413	6	Abq99374 Human cod
8	661.8	55.9	737	6	Al44090 Mouse MAR
9	651.2	55.0	2049	5	Ab574750 DNA encod
10	561.4	47.5	1348	5	Al44087 Mouse mod
11	405	34.2	603	5	Ab574748 DNA encod
12	395	33.4	445	5	Ab574747 DNA encod
13	348	29.4	444	6	Abq98669 Human ORF
14	341	28.8	875	6	Abq99151 Human ORF
15	191.2	16.2	211	5	Ab570181 DNA encod
16	158.6	13.4	2109	4	Ab502049 DNA encod
17	157.4	13.3	2665	6	Ab165189 Lung canc
18	157.4	13.3	2665	6	AbK83738 Human cDN
19	157.4	13.3	2665	7	Acc81091 Human Src
20	157.4	13.3	3756	7	Abx62975 Human act
21	141.8	12.0	432	4	Ab112879 Probe #28
22	141.8	12.0	432	4	Ab54580 Human foe
23	141.8	12.0	432	4	Ab134236 Probe #29

24	141.8	12.0	432	4	AB444128	Ab444128 Human bre
25	141.8	12.0	432	4	AB24363	Ab24363 Probe #28
26	141.8	12.0	432	4	AAK28314	AAK28314 Human bon
27	141.8	12.0	432	4	AAK02872	AAK02872 Human bra
28	141.8	12.0	432	4	AB527912	AB527912 Human liv
29	141.8	12.0	432	5	AAI02797	AAI02797 Probe #27
30	141.8	12.0	432	6	AB502823	AB502823 Human gen
31	141.8	12.0	448	4	AA114520	AA114520 Probe #44
32	141.8	12.0	448	6	AB504499	AB504499 Human gen
33	123	10.4	1926	6	ABK83940	ABK83940 Human cDN
34	123	10.4	2015	6	AB166673	AB166673 Lung canc
35	123	10.4	2015	6	ABK83939	ABK83939 Human cDN
36	123	10.4	2015	9	AD621155	AD621155 Human hae
37	123	10.4	2341	9	AD533793	AD533793 Human pro
38	119.4	10.1	2288	6	ABK83935	ABK83935 Human cDN
39	119.4	10.1	2298	9	AD519015	AD519015 Human dis
40	110.6	9.3	1530	8	ADA02968	ADA02968 Human LCK
41	110.6	9.3	1530	9	ADB72706	ADB72706 Human LCK
42	110.6	9.3	1530	9	ADC85448	ADC85448 Human LCK
43	110.6	9.3	2032	3	AA246491	AA246491 PKA subst
44	110.6	9.3	2032	7	ACC72850	ACC72850 Human can
45	110.6	9.3	2032	8	ADA02967	ADA02967 Human LCK

ALIGNMENTS

RESULT 1  
ABK61465  
ID ABK61465 standard; cDNA; 1183 BP.  
XX  
AC ABK61465;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human cDNA encoding protein NOV13.  
XX  
KW Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;  
KW cell signal processing disorder; metabolic pathway modulation disorder;  
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;  
KW uterus cancer; immune response; graft-versus-host disease;  
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;  
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;  
KW Albright hereditary osteodystrophy.  
XX  
OS Homo sapiens.  
XX  
FN WC200216599-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 27-AUG-2001; 2001WO-US026510.  
XX  
PR 25-AUG-2000; 2000US-0228191P.  
XX  
PR 08-FEB-2001; 2001US-0267300P.  
XX  
PR 20-FEB-2001; 2001US-0269961P.  
XX  
PR 20-MAR-2001; 2001US-0277337P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PA (COR-) COR THERAPEUTICS INC.  
XX  
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;  
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;  
XX  
DR WPI: 2002-280937/32.  
XX  
DR P-PSDB; AAU91308.  
XX  
PT New polypeptides for treating or preventing a disorder associated with  
XX them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.  
XX  
PS Claim 1; Page 98; 263pp; English.  
XX  
CC The invention relates to an isolated polypeptide (NOVX) a mature form of

CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide  
 CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
 CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it  
 CC and antibody against it, are useful for treating or preventing (e.g. by  
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOVX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation or Albrit  
 CC hereditary osteodystrophy and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC for developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence encodes a NOVX protein  
 XX  
 SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1183; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-310;  
 Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGCTAGAGCTCAAGGACCCAGCGCTGTCTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
 DB 1 AGCTAGAGCTCAAGGACCCAGCGCTGTCTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
 QY 61 CCTTCCCTCCCTGGCTGTGCTTGGGAGGTTCCCGAGTCCAGAACTCCCTAAGGAG 120  
 DB 61 CCTTCCCTCCCTGGCTGTGCTTGGGAGGTTCCCGAGTCCAGAACTCCCTAAGGAG 120  
 QY 121 CATGGGAGCTGATCCATCCCTGTGTACAACTGTCTGACGAGAGTGTGAGCT 180  
 DB 121 CATGGGAGCTGATCCATCCCTGTGTACAACTGTCTGACGAGAGTGTGAGCT 180  
 QY 181 ACCCAAAACCAACCTAGCTCTCCCTGAAGATCTCCCGAGCTGAGAGAGTCTGGGTG 240  
 DB 181 ACCCAAAACCAACCTAGCTCTCCCTGAAGATCTCCCGAGCTGAGAGAGTCTGGGTG 240  
 QY 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACTGTCCA 300  
 DB 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACTGTCCA 300  
 QY 301 GCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCTTTGATGACAAACCAATTCC 360  
 DB 301 GCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCTTTGATGACAAACCAATTCC 360  
 QY 361 CTCGATGATGCTCTTCTGAGTCTCTGCTGAGGAACTGCTGCCAGCAGAG 420  
 DB 361 CTCGATGATGCTCTTCTGAGTCTCTGCTGAGGAACTGCTGCCAGCAGAG 420  
 QY 421 ABAATCTGCGCAAGCCAGCTTCAAGTCTCTCTGTCAGGAGGAGGAGCTGTGACAT 480  
 DB 421 ABAATCTGCGCAAGCCAGCTTCAAGTCTCTCTGTCAGGAGGAGGAGCTGTGACAT 480  
 QY 481 GGAAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 DB 481 GGAAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 QY 541 GCGCGAGCTGTGCTGAGACTGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 DB 541 GCGCGAGCTGTGCTGAGACTGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 QY 601 GTGGACGCTGTGCTGAGGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
 DB 601 GTGGACGCTGTGCTGAGGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660

QY 661 CAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGTAGCGGGAGGAGGAGGAGGAGGAGGAG 720  
 DB 661 CAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGTAGCGGGAGGAGGAGGAGGAGGAGGAG 720  
 QY 721 GTTGTACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
 DB 721 GTTGTACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
 QY 781 CTCTTACTCTCTGTGAGTCTGGCTCAGCGGCTGTGATCTCTGAGGAGGAGGAGGAGGAG 840  
 DB 781 CTCTTACTCTCTGTGAGTCTGGCTCAGCGGCTGTGATCTCTGAGGAGGAGGAGGAGGAG 840  
 QY 841 CAGGATCCACTCTCTGTGAGTCTGGCTCAGCGGCTGTGATCTCTGAGGAGGAGGAGGAG 900  
 DB 841 CAGGATCCACTCTCTGTGAGTCTGGCTCAGCGGCTGTGATCTCTGAGGAGGAGGAGGAG 900  
 QY 901 ACTCCAGGCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 DB 901 ACTCCAGGCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 QY 961 GGAGCCCTGTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
 DB 961 GGAGCCCTGTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
 QY 1021 GACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 DB 1021 GACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 QY 1081 AGCTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
 DB 1081 AGCTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
 QY 1141 CATCAGCCTGAATGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1183  
 DB 1141 CATCAGCCTGAATGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1183  
 RESULT 2  
 AAD43980  
 ID AAD43980 standard; cDNA; 2567 BP.  
 XX AAD43980;  
 XX  
 XX AAD43980;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 DE Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.  
 XX  
 KW Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 415..1200  
 FT /\*tag= a  
 FT /product= "Human SLAP-2"  
 XX  
 XX WO200242457-A1.  
 XX  
 PN 30-MAY-2002.  
 PD  
 XX  
 PF 20-NOV-2001; 2001WO-US043367.  
 XX  
 PR 22-NOV-2000; 2000US-0252545P.  
 XX  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 XX Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;  
 PI





CC anti-HIV activities, and can be used as a modulator of lymphocyte  
CC activation, and of ubiquitination of a Cbl target protein, and in gene  
CC therapy. (i) is useful for screening a bioactive agent capable of binding  
CC to SLIM. (i) is also useful for screening a bioactive agent capable of  
CC modulating SLIM binding. (i) or its fragments is useful in the study or  
CC in the treatment of conditions which involves this function or  
CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or  
CC prevent SLIM associated disorders. (i) or the polynucleotide encoding it  
CC (ii) is useful for modulating leukocyte and/or platelet activation, for  
CC modulating antigen receptor-induced signalling and activation in  
CC leukocyte and/or platelets and for modulating antigen receptor-induced  
CC signalling and activation in lymphocytes and/or mast cells. (i) or (ii)  
CC is also useful for modulating the basal activity of lymphocytes. (i) or  
CC (ii) is useful in the treatment of immunodeficiency disorders, such as  
CC acquired immunodeficiency syndrome (AIDS), for the prevention and  
CC treatment of acute inflammatory disorders, chronic inflammatory  
CC disorders, autoimmune disorder and transplant rejection  
XX  
SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;  
Query Match 66.3%; Score 784.4; DB 6; Length 786;  
Best Local Similarity 99.9%; Pred. No. 1.9e-202;  
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGTCTCTGTC 457  
DB 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGTCTCTGTC 60  
QY 458 CAAGCCAGGACCTGTGACCATCGAGCAGAGAGAGCCAGCCAGCGGTGGCCCTG 517  
DB 61 CAAGCCAGGACCTGTGACCATCGAGCAGAGAGAGCCAGCCAGCGGTGGCCCTG 120  
QY 518 GCGAGTTTCCCGCAGGTGGCCCGCGAGCTGCTCGTAGACTCGGGAGCCATTGACC 577  
DB 121 GCGAGTTTCCCGCAGGTGGCCCGCGAGCTGCTCGTAGACTCGGGAGCCATTGACC 180  
QY 578 ATCGTCTCTGAGGATGAGAGCTGTGTGACGCTGTCTGTGAAGTCTCAGGCAGAGATAT 637  
DB 181 ATCGTCTCTGAGGATGAGAGCTGTGTGACGCTGTCTGTGAAGTCTCAGGCAGAGATAT 240  
QY 638 AACATCCCGACGCTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 697  
DB 241 AACATCCCGACGCTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 300  
QY 698 AGGAGAAAGCAGAGGAATCTCTGTGTACTCGGAAACCTGGAGGGCTTCCTCATC 757  
DB 301 AGGAGAAAGCAGAGGAATCTCTGTGTACTCGGAAACCTGGAGGGCTTCCTCATC 360  
QY 758 CGGAGAGCCAGACCCAGAGAGGCTTACTCTCTGTCTGTCAGTCCGCTCAGCCGCTGCA 817  
DB 361 CGGAGAGCCAGACCCAGAGAGGCTTACTCTCTGTCTGTCAGTCCGCTCAGCCGCTGCA 420  
QY 818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCATATGGCTGGCTGTACATC 877  
DB 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCATATGGCTGGCTGTACATC 480  
QY 878 TCACCGGCTTCACTTCCCTTCACTCAGGCCCTGTGTGACCACTTACTCTGAGCTGGCG 937  
DB 481 TCACCGGCTTCACTTCCCTTCACTCAGGCCCTGTGTGACCACTTACTCTGAGCTGGCG 540  
QY 938 GATGACATCTGTCTACTCAAGAGGCTGTGTCTCTGAGAGGCTGGCCGCTCCCT 997  
DB 541 GATGACATCTGTCTACTCAAGAGGCTGTGTCTCTGAGAGGCTGGCCGCTCCCT 600  
QY 998 GGCAGAGGATATACCCCTTACCTGTGTGTGAGAGGACACCCTCACTCACTGAAGAGCTG 1057  
DB 601 GGCAGAGGATATACCCCTTACCTGTGTGTGAGAGGACACCCTCACTCACTGAAGAGCTG 660  
QY 1058 GACAGCTCCCTCTCTTTCTGAGAGCTGCCACAGGGAGAGGTCTCTTCTCAGTGAGGGT 1117  
DB 661 GACAGCTCCCTCTCTTTCTGAGAGCTGCCACAGGGAGAGGTCTCTTCTCAGTGAGGGT 720  
QY 1118 CTCGGGAGTCCCTCAGCTTCTATCATCAGCCTGAAATGACGAGGCTGTCTCTTTGGATGAT 1177

DB 721 CTCGGGAGTCCCTCAGCTTCTATCATCAGCTGATGAGGCTGTCTCTTTGGATGAT 780  
QY 1178 GCTTAG 1183  
DB 781 GCTTAG 786  
RESULT 5  
ABQ98670  
ID ABQ98670 standard; DNA; 763 BP.  
XX  
AC ABQ98670;  
XX  
DT 04-NOV-2002 (first entry)  
XX  
DE Human ORF477 coding sequence.  
XX  
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulneryary;  
KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
KW cancer; cardiovascular disease; allergy; autoimmune disease;  
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
XX  
OS Homo sapiens.  
XX  
PN US2002082206-A1.  
XX  
PD 27-JUN-2002.  
XX  
PF 30-MAY-2001; 2001US-00867550.  
XX  
PR 30-MAY-2000; 2000US-0208427P.  
XX  
PA (LEAC/) LEACH M D.  
PA (MEHR/) MEHRABAN F.  
PA (CONL/) CONLEY P B.  
PA (TOPP/) TOPPER J N.  
PA (LAWD/) LAW D.  
XX  
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
XX  
WI; 2002-626554/57.  
DR P-PSDB; ABP64107.  
XX  
PT New polypeptide designated ORFX are present in human atherogenic cells  
PT and are useful to prevent and treat ORFX-associated disorders including  
PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
PT inflammatory disease.  
XX  
PS Claim 2; SEQ ID NO 953; 78pp; English.  
XX  
CC The present invention relates to novel human ORFX polypeptides and their  
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
CC were discovered in human atherogenic cells, in particular in platelets  
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
CC many other tissues as well. Atherogenic cells are cells which have the  
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
CC nucleic acids are useful for treating or preventing a pathological  
CC condition associated with an ORFX-associated disorder, e.g. cancer,  
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
CC coagulation disorders or inflammatory disorders. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from the USPTO web site at  
CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
XX  
SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;  
Query Match 64.1%; Score 758.2; DB 6; Length 763;  
Best Local Similarity 99.6%; Pred. No. 2.4e-195;  
Matches 760; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 113 CTAAGAGCATGGGACGCTGATCCATCCCTGGTGTACAACTGCTGACTGACGACAGAT 172

Db 1 CTATGAGCATGGGAGAGTGTATCCATCCCTGCTGTATCAAACTGTGATCGCAGACAGAT 60  
 QY 173 GCTGAGCTACCCAAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAT 232  
 Db 61 GCTGAGCTACCCAAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAT 120  
 QY 233 TCTGGTGTCTAGGACCAAGGACATGGGAGACTTCCAGAGGGCCCCAAAGCCCTAA 292  
 Db 121 TCTGGATGTCTAGGACCAAGGACATGGGAGACTTCCAGAGGGCCCCAAAGCCCTAA 180  
 QY 293 CTTGTCCAGCCAGAGCATGGTCTCAGCAGAGCTGCTTCCCAAGGCTTTGATGACAAAC 352  
 Db 181 CTTGTCCAGCCAGAGCATGGTCTCAGCAGAGCTGCTTCCCAAGGCTTTGATGACAAAC 240  
 QY 353 CAATTTCCCTCGATGTGCTTCTGAGTCTCTGCTGAGGACCAATGGGAAGTCTGCC 412  
 Db 241 CAATTTCCCTCGATGTGCTTCTGAGTCTCTGCTGAGGACCAATGGGAAGTCTGCC 300  
 QY 413 AGCAGAGAAATCTCTGCCAAGCCCAAGCTTCAAGTTCCTCTGTCCAAAGCCAGGACCT 472  
 Db 301 AGCAGAGAAATCTCTGCCAAGCCCAAGCTTCAAGTTCCTCTGTCCAAAGCCAGGACCT 360  
 QY 473 GTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGCTGCTGGGAGTTTCCCGGCA 532  
 Db 361 GTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGCTGCTGGGAGTTTCCCGGCA 420  
 QY 533 GGTGGCCGCGAGCTGCTGCTGAGACTCGGGAGCCATGACCATCTGCTCTGAGGAT 592  
 Db 421 GGTGGCCGCGAGCTGCTGCTGAGACTCGGGAGCCATGACCATCTGCTCTGAGGAT 480  
 QY 593 GGAGACTGTGGAGCGGTGCTGCTGAGTCTCAGGAGAGAGTATTAACATCCCGCGTC 652  
 Db 481 GGAGACTGTGGAGCGGTGCTGCTGAGTCTCAGGAGAGAGTATTAACATCCCGCGTC 540  
 QY 653 CAGTGGCCAAAGTCTCCATGGTGGCTGTATGAGGCGCTGAGCAGGAGAGAAAGCAG 712  
 Db 541 CAGTGGCCAAAGTCTCCATGGTGGCTGTATGAGGCGCTGAGCAGGAGAGAAAGCAG 600  
 QY 713 GAACCTGCTGTGTTACTCTGGAAACCTCTGAGGGGCTTCTCATCCGGGAGAGCCAGAC 772  
 Db 601 GAACCTGCTGTGTTACTCTGGAAACCTCTGAGGGGCTTCTCATCCGGGAGAGCCAGAC 660  
 QY 773 AGAGAGGCTCTACTCTCTGATCGGCTCAGCGGCTGATGAGGAGAGAGAGAGAGAG 832  
 Db 661 AGAGAGGCTCTACTCTCTGATCGGCTCAGCGGCTGATGAGGAGAGAGAGAGAGAG 720  
 QY 833 AGACACTACAGGATCCACTCCCTTGACAATGGCTGGTGTACA 875  
 Db 721 AGACACTACAGGATCCACTCCCTTGACAATGGCTGGTGTACA 763

RESULT 6

AAC77202  
 ID AAC77202 standard; cDNA; 837 BP.

AC AAC77202;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2757 polynucleotide sequence SEQ ID NO:5513.

Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 vulnerability; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;  
 anticonvulsant; osteopathic; aniaarthritic; immunosuppressant; cardiac;  
 immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 hypotensive; dermatological; immunosuppressive; antinflammatory;  
 antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 neurodegenerative disease; osteoarthritis; graft vs host disease;  
 cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 cholesterol ester storage; systemic lupus erythematosus; infection;  
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.  
 OS Homo sapiens.  
 XX WO200058473-A2.  
 PN 05-OCT-2000.  
 XX 31-MAR-2000; 2000WO-US008621.  
 XX 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Leach M;  
 XX WPI; 2000-602362/57.  
 DR P-PSDB; AAB42993.  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX Claim 5; Page 4692-4693; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 CC antipsoriatic; antiparkinsonian; neurotropic; neuroprotective; osteopathic;  
 CC anticonvulsant; aniaarthritic; immunosuppressant; immunostimulant;  
 CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antinflammatory; antibacterial; The  
 CC antiviral; antirheumatic; antithyroid; and antianemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmunity  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX Sequence 837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;

Query Match 62.6%; Score 740.2; DB 3; Length 837;

Best Local Similarity 99.6%; Pred. No. 1.9e-190;

Matches 742; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 439 AAGCTTGAGTTCCTCTGTCGCAAGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAA 498  
 Db 3 AAGCTTGAGTTCCTCTGTCGCAAGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAA 62  
 QY 499 GGCCACAGCCGCGCCCTGCGGACGTTTCCCGGACAGGTGGCCCGCGAGTGTGCTGAG 558  
 Db 63 GGCCACAGCCGCGCCCTGCGGACGTTTCCCGGACAGGTGGCCCGCGAGTGTGCTGAG 122  
 QY 559 ACTGGGGGAGCATTGACCATCTGCTGAGGATGAGAGCTGGGAGCGGTGCTGCTGA 618  
 Db 123 ACTGGGGGAGCATTGACCATCTGCTGAGGATGAGAGCTGGGAGCGGTGCTGCTGA 182  
 QY 619 AGTCTCAGGAGAGATATTAACATCCCGAGCGTCCACGTGGGCAAGTCTCCCATGGGTG 678  
 Db 183 AGTCTCAGGAGAGATATTAACATCCCGAGCGTCCACGTGGGCAAGTCTCCCATGGGTG 242

XX New isolated polynucleotide, useful in research, diagnostic or  
PT therapeutic methods, e.g. preventing or treating disorders involving  
PT aberrant protein expression or biological activity.  
XX  
XX  
XX Claim 1: SEQ ID NO 107; 394pp; English.  
XX  
XX The present invention relates to novel human coding sequences (AB099268-  
CC AB099608) and proteins (AB964382-ABP5022). The sequences are useful in  
CC therapeutic, diagnostic and research methods. The polynucleotides may be  
CC used in the field of molecular biology as hybridisation probes, primers  
CC for PCR, for chromosome and gene mapping, for the recombinant production  
CC of protein, or in generation of anti-sense DNA or RNA. The  
CC polynucleotides are useful in diagnostics as expressed sequence tags  
CC (ESTs) for identifying expressed genes or for physical mapping of the  
CC human genome. The proteins may be used as molecular weight markers, or as  
CC nutritional sources or supplements. The proteins may be used to maintain  
CC and expand cell population in a totipotent or pluripotential state  
CC useful for re-engineering damaged or diseased tissues, transplantation,  
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
CC polynucleotides and proteins are useful for preventing, treating or  
CC ameliorating disorders involving aberrant protein expression or  
CC biological activity, e.g. haematopoietic disorders, central/peripheral  
CC nervous system diseases, mechanical and traumatic disorders, non-healing  
CC wounds, immune deficiencies and disorders, infectious diseases caused by  
CC viral, bacterial or fungal infection, autoimmune disorders, allergic  
CC reactions and conditions, coagulation disorders, or cancer. The  
CC polynucleotide sequences of the invention were assembled from ESTs  
CC isolated mainly by sequencing by hybridisation, and in some cases,  
CC sequences obtained from one or more public databases. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIFO at  
CC ftp.wifo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1413 BP: 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other

Query Match	61.2%;	Score 724.2;	DB 6;	Length 1413;
-------------	--------	--------------	-------	--------------

```
Best Local Similarity 93.7%; Pred. No. 4.9e-186;
Matches 786; Conservative 0; Mismatches 3;
Indels 50; Gaps 1
```

[illegible]

345 TGACAAACCAATTTCCTCGATGATGTGCTTCTGAGTGTCTCTGCTGAGGAACAAATGGGA 401  
QY

Db  
1 TGACAAACCAATTCCCTCGATGATGTGCTTCTGAGTCTCTGCTGAGGAACAATGGAG 60

405 GTCTGCCCGCAGAGAAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCTCCAAAGGCC 464

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

DB 61 GTC TGG CCG CAG CAG AAG AAA T C T C T G C C A A G C C C A A G C T T G A G T T C C T C T G T C C A A G G C C 120

QY 465 AGGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTGGGCAGTT 524

Db  
121 AGGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTGGGCAGTT 180

[illegible]

525 TCCGGGAGGTGGCCCGGGCCGAGCTGTCCCTGAGACTCGGGAGCTCATTGACCATCTGCTCT  
526

Db 181 TCCCGGAGGTGGCCCGCCGAGCTGTCGCTGAGACTCGGGAGCCATTGACCATCGTCT 240

QY 585 CTGAGGATGGAGACTGGTGGACGGTGTCTGAAGTCTCAGGCAGAGAGTATAACATCC 644

[illegible]

DD 241 C1GAGGATGGAGAC1GG1GGACGG1GC1GTC1GAAATC1CAGGCGAGAGATATACATCC 300

Qy  
645 CCAGCGTCCACGTGGGCAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGCAGGGAGA 704  
Qy

Db 301 CCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGCAGGGAGA 360

705 AAGCAGAGGAACTGCTGTTGTTACCTGGGAACCTGGAGGGGCTTCTCTCATCCGGGAGA 764

[illegible]

Db 361 AAGCAGAGGAACTGCTGTTGTTACCTGGGAACCCCTGGAGGGGCCTTCCTCATCCGGGAGA 420

QY 765 GCCAGACCAGGAGAGGCTCTTACTCTCTGTCTAGTCCGCCCTCAGCCGCCCTGCATCCTGGG 824

Db 421 GCCAGACCGGAGAGGCTTTACTCTCTGTCTCAGTCCGCCCTCAGCCGCCCTGCATCCCTGG 480

[illegible]

QY  
825 ACCGGATCAGACACTACAGGATCCACTGCCCTTGACAAATGGCTGGCTGTACATCTCACCGC 881









CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;  
Query Match 34.2%; Score 405; DB 5; Length 603;  
Best Local Similarity 100.0%; Pred. No. 1.4e-99;  
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 779 GGCCTTACTCTCTGTCAGTCGCTCAGCCGCTGATCTGGACCGGATCAGACAC 838  
DB 199 GGCCTTACTCTCTGTCAGTCGCTCAGCCGCTGATCTGGACCGGATCAGACAC 258  
QY 839 TACAGATCCACTGGCTTGACATGGCTGGCTGATCTACATCTACCGCGCTCACCTTCCC 898  
DB 259 TACAGATCCACTGGCTTGACATGGCTGGCTGATCTACATCTACCGCGCTCACCTTCCC 318  
QY 899 TCATCCAGCCGCTGGTGACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTC 958  
DB 319 TCATCCAGCCGCTGGTGACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTC 378  
QY 959 AAGGAGCCCTGTCTCTGTCAGAGGCTGGCGCTTCCCTGCGAGGATATACCCCTACCT 1018  
DB 379 AAGGAGCCCTGTCTCTGTCAGAGGCTGGCGCTTCCCTGCGAGGATATACCCCTACCT 438  
QY 1019 GTGACTGTGTCAGAGGACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCT 1078  
DB 439 GTGACTGTGTCAGAGGACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCT 498  
QY 1079 GAAGTGTCCACAGGGAGGAGTCTTCTCTAGTGGGTCTCCGGAGTCCCTCAGCTTC 1138  
DB 499 GAAGTGTCCACAGGGAGGAGTCTTCTCTAGTGGGTCTCCGGAGTCCCTCAGCTTC 558  
QY 1139 TACATCAGCTGAATGACGAGGCTCTCTTTGGATGATGCTAG 1183  
DB 559 TACATCAGCTGAATGACGAGGCTCTCTTTGGATGATGCTAG 603

RESULT 12  
AAS74747/c  
ID AAS74747 standard; cDNA; 445 BP.  
XX AAS74747;  
AC  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #10551.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.

DR P-PSDB; ABG10560.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1; SEQ ID NO 10551; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC reactions. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 445 BP; 89 A; 112 C; 143 G; 101 T; 0 U; 0 Other;  
Query Match 33.4%; Score 395; DB 5; Length 445;  
Best Local Similarity 98.8%; Pred. No. 6.5e-97;  
Matches 409; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
QY 54 CCCTGGGCTTCCCTCCCTGGCTCGGCTGTGTTGGAGGGTTCGCCAGTCAGAAATCCC 113  
DB 416 CGCTGGGCTTCCCTCCCTGGCTCGGCTGTGTTGGAGGGTTCGCCAGTCAGAAATCCC 357  
QY 114 TAAGGAGCATGGGCGAGCTGATCCATCCCTGGTGTACAACTGCTGCTGACGACAGATG 173  
DB 356 TAAGGAGCATGGGCGAGCTGATCCATCCCTGGTGTACAACTGCTGCTGACGACAGATG 297  
QY 174 CTGAGCTAGCCAAACCAACACCTAGCTCTCCCTGAAGATCCCTCCAGCTGAGAGATG - 232  
DB 296 CTGAGCTAGCCAAACCAACACCTAGCTCTCCCTGAAGATCCCTCCAGCTGAGAGATG 237  
QY 233 TCTGGGTGCTTAGGACCAAGGACACTGGCAGACTTCAGAGGGGCCCCCAAGCCCTAA 292  
DB 236 TCTGGGTGCTTAGGACCAAGGACACTGGCAGACTTCAGAGGGGCCCCCAAGCCCTAA 177  
QY 293 CTTGTCCAGCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCTTTGATGACAAAC 352  
DB 176 CTTGTCCAGCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCTTTGATGACAAAC 117  
QY 353 CAATTTCCCTCGATGATGTGCTTCTGAGTGCTCTGCTGAGGAAACAATGGAGTCTGCC 412  
DB 116 CAATTTCCCTCGATGATGTGCTTCTGAGTGCTCTGCTGAGGAAACAATGGAGTCTGCC 57  
QY 413 AGCAGAGGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCCCAAGCCAGG 467  
DB 56 AGCAGAGGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCCCAAGCCAGG 2

RESULT 13  
ABQ98669  
ID ABQ98669 standard; DNA; 444 BP.  
XX  
XX ABQ98669;  
AC  
XX  
XX 04-NOV-2002 (first entry)  
DT

XX DE Human ORF476 coding sequence.  
 XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;  
 XX KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 XX KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 XX KW cancer; cardiovascular disease; allergy; autoimmune disease; ds.  
 XX KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 XX OS Homo sapiens.  
 XX PN US2002082206-A1.  
 XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX PR 30-MAY-2000; 2000US-0208427P.  
 XX PA (LEAC/) LEACH M D.  
 XX PA (MEHR/) MEHRABAN F.  
 XX PA (CONL/) CONLEY P B.  
 XX PA (TOPP/) TOPPER J N.  
 XX PA (LAWD/) LAW D.  
 XX FI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX DR WPI; 2002-626554/67.  
 XX DR P-PSDB; ABP64106.  
 XX PT New polypeptide designated ORFX are present in human atherogenic cells  
 XX PT and are useful to prevent and treat ORFX-associated disorders including  
 XX PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 XX PT inflammatory disease.  
 XX PS Claim 2; SEQ ID NO 951; 78pp; English.  
 XX CC The present invention relates to novel human ORFX polypeptides and their  
 XX CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 XX CC were discovered in human atherogenic cells, in particular in platelets  
 XX CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 XX CC many other tissues as well. Atherogenic cells are cells which have the  
 XX CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 XX CC nucleic acids are useful for treating or preventing a pathological  
 XX CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 XX CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 XX CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 XX CC for this patent did not form part of the printed specification, but was  
 XX CC obtained in electronic format directly from the USPTO web site at  
 XX CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
 XX SQ Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;  
 XX CC Query Match 29.4%; Score 348; DB 6; Length 444;  
 XX CC Best Local Similarity 100.0%; Pred. No. 3.7e-84;  
 XX CC Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 242 CCTAGGACCAAGACATGCGAGCTTCCAGAGGGCCCCCAAGCCCTAACCTGTCAG 301  
 DB 1 CCTAGGACCAAGACATGCGAGCTTCCAGAGGGCCCCCAAGCCCTAACCTGTCAG 60  
 QY 302 CCAGAGCATGCGTCTCAGCAGAGCTGTCTCCCAAGCCTTTGATGACAAACCAATTTC 361  
 DB 61 CCAGAGCATGCGTCTCAGCAGAGCTGTCTCCCAAGCCTTTGATGACAAACCAATTTC 120  
 QY 362 TCGATGATGCTCTTCTGAGTCTCTGCTGAGGACATGGAGTCTGCCACACAGAGA 421  
 DB 121 TCGATGATGCTCTTCTGAGTCTCTGCTGAGGACATGGAGTCTGCCACACAGAGA 180  
 QY 422 AAATCTCTGCCAAGCCCAAGCTTCTCTCTCTCCAAAGCCAGGACCTGTGACCATG 481  
 DB 181 AAATCTCTGCCAAGCCCAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240

QY 482 GAAGCAGAGAGAAAGCAAGCCGACAGCCCTGGCCCTCTGGGAGTTTCCCGCAGGTGGCCCG 541  
 DB 241 GAAGCAGAGAGAAAGCAAGCCGACAGCCCTGGCCCTCTGGGAGTTTCCCGCAGGTGGCCCG 300  
 QY 542 GCCGAGCTGTGCTGAGACTCGGGAGGACATGACCATCGTCTCTGAG 589  
 DB 301 GCCGAGCTGTGCTGAGACTCGGGAGGACATGACCATCGTCTCTGAG 348  
 RESULT 14  
 ABQ99151  
 ID ABQ99151 standard; DNA; 875 BP.  
 XX AC ABQ99151;  
 XX DT 04-NOV-2002 (first entry)  
 XX DE Human ORF958 coding sequence.  
 XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;  
 XX KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 XX KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 XX KW cancer; cardiovascular disease; allergy; autoimmune disease; ds.  
 XX KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 XX OS Homo sapiens.  
 XX PN US2002082206-A1.  
 XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX PR 30-MAY-2000; 2000US-0208427P.  
 XX PA (LEAC/) LEACH M D.  
 XX PA (MEHR/) MEHRABAN F.  
 XX PA (CONL/) CONLEY P B.  
 XX PA (TOPP/) TOPPER J N.  
 XX PA (LAWD/) LAW D.  
 XX FI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX DR WPI; 2002-626554/67.  
 XX DR P-PSDB; ABP64588.  
 XX PT New polypeptide designated ORFX are present in human atherogenic cells  
 XX PT and are useful to prevent and treat ORFX-associated disorders including  
 XX PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 XX PT inflammatory disease.  
 XX PS Claim 2; SEQ ID NO 1915; 78pp; English.  
 XX CC The present invention relates to novel human ORFX polypeptides and their  
 XX CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 XX CC were discovered in human atherogenic cells, in particular in platelets  
 XX CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 XX CC many other tissues as well. Atherogenic cells are cells which have the  
 XX CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 XX CC nucleic acids are useful for treating or preventing a pathological  
 XX CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 XX CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 XX CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 XX CC for this patent did not form part of the printed specification, but was  
 XX CC obtained in electronic format directly from the USPTO web site at  
 XX CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
 XX SQ Sequence 875 BP; 205 A; 259 C; 225 G; 185 T; 0 U; 1 Other;  
 XX CC Query Match 28.8%; Score 341; DB 6; Length 875;  
 XX CC Best Local Similarity 100.0%; Pred. No. 3.7e-82;  
 XX CC Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 843 GGATCCACTGCTTGCATGATGCTGCTGATCATCTCACCGGCGCTCACCTTCCCTCAC 902  
Db 2 GGATCCACTGCTTGCATGATGCTGCTGATCATCTCACCGGCGCTCACCTTCCCTCAC 61  
QY 903 TCCAGGCCCTGGTGGACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTTACTCAAGG 962  
Db 62 TCCAGGCCCTGGTGGACCACTTACTCTGAGCTGGCGGATGACATCTGCTTACTCAAGG 121  
QY 963 AGCCCTGTCTCTGCAGAGGGTGGCCCGCTCCCTGGGAGGATATACCTTACTGTGA 1022  
Db 122 AGCCCTGTCTCTGCAGAGGGTGGCCCGCTCCCTGGGAGGATATACCTTACTGTGA 181  
QY 1023 CTGTGCAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAAG 1082  
Db 182 CTGTGCAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAAG 241  
QY 1083 CTGCCACAGGGAGGAGTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACA 1142  
Db 242 CTGCCACAGGGAGGAGTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACA 301  
QY 1143 TCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCCTAG 1183  
Db 302 TCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCCTAG 342

RESULT 15

AAS70181  
ID AAS70181 standard; cDNA; 211 BP.

AC AAS70181;

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #5985.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-539362/73.

XX P-PSDB; ABG05994.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 1; SEQ ID NO 5985; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 211 BP; 50 A; 51 C; 72 G; 38 T; 0 U; 0 Other;

Query Match 16.2%; Score 191.2; DB 5; Length 211;  
Best Local Similarity 98.3%; Pred. No. 9.6e-42;  
Matches 193; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 584 TCTGAGGATGGAGCTGGTGGACGGTGTCTCTGAAAGTCTCAGGCAGAGATATAATC 643  
Db 16 TCCAAGGATGGAGCTGGTGGACGGTGTCTCTGAAAGTCTCAGGCAGAGATATAATC 75  
Qy 644 CCCAGGCTCCACGTGGGCAAGTCTCCATGGGTGCTGTATGAGGCTGAGCAGGGAG 703  
Db 76 CCCAGGCTCCACGTGGGCAAGTCTCCATGGGTGCTGTATGAGGCTGAGCAGGGAG 135  
Qy 704 AAAGCAGAGGAACCTGCTGTTGTTTACCTGGAAACCTGGAGGGGCTTCTCTCATCCGGAG 763  
Db 136 AAAGCAGAGGAACCTGCTGTTGTTTACCTGGAAACCTGGAGGGGCTTCTCTCATCCGGAG 195  
Qy 764 AGCCAGACCAAGGAGAG 779  
Db 196 AGCCAGACCAAGGAGAG 211

Search completed: February 19, 2004, 21:51:39  
Job time : 1105 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:54 ; Search time 7832.77 Seconds  
(without alignments)  
4510.152 Million cell updates/sec

Title: US-09-939-853A-74  
Perfect score: 1183  
Sequence: 1 agctagagctcaaggacc.....tctcttgatgagcctag 1183

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_hc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_hc.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pln.\*
- 20: em\_gss\_vrt.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_nam.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_phg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gss1.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	921.4	77.9	1201	9 AL541041	AL541041 AL541041
2	860.2	72.7	1020	12 BQ054281	BQ054281 AGENCOURT
3	817	69.1	1002	12 BQ052308	BQ052308 AGENCOURT
4	731.4	61.8	1069	12 BQ052468	BQ052468 AGENCOURT

5	673.6	56.9	778	12	BG178487
6	662.6	56.0	986	12	BQ054265
7	589	49.8	878	12	BQ053486
8	561.4	47.5	2637	11	AK088672
9	546.8	46.2	616	13	BX383606
10	538.6	45.5	2974	11	AK030877
11	476.4	40.3	296	11	AK020837
12	452.2	38.2	660	13	BY742155
13	405	27.3	566	12	BG284179
14	322.4	24.3	537	9	AL844311
15	321	27.1	960	13	BU944126
16	319.8	27.0	660	10	BG635615
17	317.4	26.8	614	9	AL844307
18	314.6	26.6	606	9	AL844309
19	282	23.8	377	9	AA959151
20	278	23.5	627	10	BB619854
21	277.6	23.5	569	14	CB426333
22	272.8	23.1	541	12	BI989899
23	261.4	22.1	794	12	BG677567
24	243.6	20.6	670	9	AI510095
25	239.4	20.2	640	13	BQ553006
26	233.6	19.7	389	10	BE015229
27	221.6	18.7	603	13	BQ553005
28	221.4	18.7	439	10	BE015226
29	174	14.7	701	13	BX849096
30	169.2	14.3	569	14	CB514688
31	166.8	14.1	642	9	AL844308
32	161.8	13.7	619	9	AL844312
33	161.8	13.7	831	29	AY416277
34	160.6	13.6	871	13	BQ436143
35	159	13.4	570	9	AI471720
36	158.6	13.4	1003	13	BX415149
37	158.6	13.4	1106	13	BX436423
38	158.6	13.4	1133	9	AL549826
39	158.6	13.4	1201	9	AL539427
40	158.2	13.4	1201	9	AL551370
41	157.6	13.3	1063	13	BQ072745
42	157.4	13.3	972	13	BQ707614
43	150.2	12.7	849	12	BI769183
44	149.8	12.7	653	13	BY746881
45	138.4	11.7	611	9	AL844310

## ALIGNMENTS

RESULT 1	AL541041	AL541041	1201 bp	mRNA	linear	EST 12-MAY-2003
LOCUS	AL541041	Homo sapiens PLACENTA	Homo sapiens	cdna clone	CS00005YK23	
DEFINITION	5-PRIME, mRNA	sequence.				
ACCESSION	AL541041					
VERSION	AL541041.2	GI:30544829				
KEYWORDS	EST.					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 1201)					
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.					
TITLE	Full-length cDNA libraries and normalization					
JOURNAL	Unpublished (2001)					
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12871733.					
	Contact: Genoscope					
	Genoscope - Centre National de Sequencage					
	BP 131 91006 EVRY cedex - France					
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr					
	Library was constructed by Life Technologies, a division of					
	Invitrogen. This sequence belongs to sequence cluster 9825.r For					
	more information about this cluster, see					
	http://www.genoscope.cns.fr/					
	csi-bin/cluster.cgi?seq=CS00005AF12QPl&cluster=9825.r. Contact :					
	Feng Liang Email : fliang@lifetech.com URL :					

```

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005AF120F1.
FEATURES
    source              1..1201
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="CS0DE005YK23"
                        /tissue_type="PLACENTA"
                        /clone_lib="Homo sapiens PLACENTA"
                        /notes="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                        with a NotI-oligo(dT) primer. Five prime end enriched,
                        double-strand cDNA was digested with Not I and cloned into
                        the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                        Library was not normalized."
ORIGIN
Query Match      77.9%; Score 921.4; DB 9; Length 1201;
Best Local Similarity 97.9%; Pred. No. 1.5e-221;
Matches 919; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY      4   TAGAGTCCAAAGACCCACGCTGTGTCTCTGTGACAGAGCTCAAGGGCCCTGGGCGCT 63
Db      65   TAGAGTCCAAAGACCCACGCTGTGTCTCTGTGACAGAGCTCAAGGGCCCTGGGCGCT 124

QY      64   TCCTCCCTGGCTCGGTGTGCTTGGAGGGTTCCCCAGTCCAGATCCCTAAGGAGCAT 123
Db      125  TCCTCCCTGGCTCGGTGTGCTTGGAGGGTTCCCCAGTCCAGATCCCTAAGGAGCAT 184

QY      124  GGGGAGCTGATCATCCCTCGGTGTACAACTGCTGACTGCAGACGATGCTGAGTACC 183
Db      185  GGGGAGCTGATCATCCCTCGGTGTACAACTGCTGACTGCAGACGATGCTGAGTACC 244

QY      184  CAAACCAACACCTAGCCTCTCCCTGAAGATCCTCCAGGCTGAGAGAGTTCTGGGTGTCC 243
Db      245  CAAACCAACACCTAGCCTCTCCCTGAAGATCCTCCAGGCTGAGAGAGTTCTGGGTGTCC 304

QY      244  TAGACCAAGACACTGGGAGACTTCAGAAAGGGCCCCAAGCCCTAACTGTCGAGCC 303
Db      305  TAGACCAAGACACTGGGAGACTTCAGAAAGGGCCCCAAGCCCTAACTGTCGAGCC 364

QY      304  AGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGATGACAAACCAATTTCCCTC 363
Db      365  AGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGATGACAAACCAATTTCCCTC 424

QY      364  GATGATGTGCTTCTGAGTGTCTCTGAGGAACAATGGGAAGTCTGCCAGCAGAGAAA 423
Db      425  GATGATGTGCTTCTGAGTGTCTCTGAGGAACAATGGGAAGTCTGCCAGCAGAGAAA 484

QY      424  ATCTCTCCCAAGCCCAAGCTTGATTCTCTGTCCAAAGGCCAGGGAACCTGTGACCATGGA 483
Db      485  ATCTCTCCCAAGCCCAAGCTTGATTCTCTGTCCAAAGGCCAGGGAACCTGTGACCATGGA 544

QY      484  AGCAGAGAGAAGCAGGCGCAGCCGTGGCCCTGGGAGTTTCCCGCAGGTGGCCCGCC 543
Db      545  AGCAGAGAGAAGCAGGCGCAGCCGTGGCCCTGGGAGTTTCCCGCAGGTGGCCCGCC 604

QY      544  CGAGCTGTGCTGAGACTCGGGGAGCCATTGACCATCGTCTCTGAGATGGAGACTGGTG 603
Db      605  CGAGCTGTGCTGAGACTCGGGGAGCCATTGACCATCGTCTCTGAGATGGAGACTGGTG 664

QY      604  GACGGTGTCTTGAAGTCTCAGCAGAGAGTATPAAATCCCCAGCCTTCCAGTGGGCAA 663
Db      665  GACGGTGTCTTGAAGTCTCAGCAGAGAGTATPAAATCCCCAGCCTTCCAGTGGGCAA 724

QY      664  AGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGCAGGGAGAGAAACAGAGAACTGCTGTT 723
Db      725  AGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGCAGGGAGAGAAACAGAGAACTGCTGTT 784

QY      724  GTTACCTGGGAAACCTTGGAGGGGCTTCTCTCATCCGGAGAGCCAGACCGAGAGGCTC 783
Db      785  GTTACCTGGGAAACCTTGGAGGGGCTTCTCTCATCCGGAGAGCCAGACCGAGAGGCTC 844

```



Db 121 GCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGTATGACAAACCAATTTCCCTCGAT 180  
Qy 367 CATGTCCTTCTGAGTGTCTCTCTGAGGAACAATGGGAAGTCTGCCAGCAGAGAAATC 426  
Db 181 GATGTCCTTCTGAGTGTCTCTCTGAGGACATATGGGAAGTCTGCCAGCAGAGAAATC 240  
Qy 427 TCTGCCAAGCCCAAGCTTGAAGTCTCTCTGTCACCAAGGCCAGGACCTGTGACCATGGAAGC 486  
Db 241 TCTGCCAAGCCCAAGCTTGAAGTCTCTCTGTCACCAAGGCCAGGACCTGTGACCATGGAAGC 300  
Qy 487 AGAGAGAAGCAAGGCCACAGCGTGGCCCTGGGAGTTCCTCCGCGAGTGGCCGCGCGA 546  
Db 301 AGAGAGAAGCAAGGCCACAGCGTGGCCCTGGGAGTTCCTCCGCGAGTGGCCGCGCGA 360  
Qy 547 GCTGTGCTGAGACTCGGGGAGCCATTGACCATCTCTGAGGATGAGACTGTTGGAC 606  
Db 361 GCTGTGCTGAGACTCGGGGAGCCATTGACCATCTCTGAGGATGAGACTGTTGGAC 420  
Qy 607 GGTGCTGTCAAGTCTCAGGACAGAGATTAACATCCCGAGTGCACCTGGGCAAGT 666  
Db 421 GGTGCTGTCAAGTCTCAGGACAGAGATTAACATCCCGAGTGCACCTGGGCAAGT 480  
Qy 667 CTCCATGGTGGTGTATGAGGCGCTGAGCAGGAGAAAGCAGAGAACTGCTGTTGTT 726  
Db 481 CTCCATGGTGGTGTATGAGGCGCTGAGCAGGAGAAAGCAGAGAACTGCTGTTGTT 540  
Qy 727 ACCTGGGAACCTCGAGGGGCTTCTCTATCCGGAGAGCCAGACAGAGAGGCTTTA 786  
Db 541 ACCTGGGAACCTCGAGGGGCTTCTCTATCCGGAGAGCCAGACAGAGAGGCTTTA 600  
Qy 787 CTCTCTGTGAGTCCGCTCAGCGCCCTGATCTGGGACCGGATCAGACACTACAGGAT 846  
Db 601 CTCTCTGTGAGTCCGCTCAGCGCCCTGATCTGGGACCGGATCAGACACTACAGGAT 660  
Qy 847 CCATGCTTGACAAATGGTGGTGTATGAGGCGCTGAGCAGGAGAAAGCAGAGAACTGCTGTTGTT 906  
Db 661 CCATGCTTGACAAATGGTGGTGTATGAGGCGCTGAGCAGGAGAAAGCAGAGAACTGCTGTTGTT 720  
Qy 907 GG-CCCTGGTGGACCAATTACTCTGAGTGGGGATGACATCTGTGCTA-CTCAAGGAG 964  
Db 721 GGCCCTGGTGGACCAATTACTCTGAGTGGGGATGACATCTGTGCTA-CTCAAGGAG 780  
Qy 965 CCCTGTGCTGAGAGGGTGGCCGCTCCTGGCAAGGATATA-CCCTTACTGTGAC 1023  
Db 781 CCCTGTGCTGCAAGGGGTGGCCGCTCCTGGCAAGGATATA-CCCTTACTGTGAC 840  
Qy 1024 TGTCAGAGGACCACTCACTGAGAGAGTGGCCGCTCCTGGCAAGGATATA-CCCTTACTGTGAC 1081  
Db 841 TGTCAGAGGACCACTCACTGAGAGAGTGGCCGCTCCTGGCAAGGATATA-CCCTTACTGTGAC 900  
Qy 1082 GCTGCCACA-GGGAGAGTCTCTCTCAGTGAGG--TCTCCGGAGTCCCTCAGCTTC 1138  
Db 901 GCTGCCCCAGGGGAGAGTCTCTCTCTCATTGAGGGGTCTTCGGGAGTCCCTCAGCTTC 960  
Qy 1139 TACATCAGCTG-NATGACGAGGTGTCTCTTTGG 1172  
Db 961 TACATCAGCTGTAATGACAGGGTGGCCCTTTG 995

RESULT 3  
LOCUS BQ052308  
DEFINITION AGENCOURT\_686571 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5933542  
5', mRNA sequence.  
ACCESSION BQ052308  
VERSION BQ052308.1 GI:19811648  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1002)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2118 row: d column: 23  
High quality sequence stop: 670.

FEATURES  
source  
1..1002  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5933542"  
/tissue\_type="natural killer cells, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 106"  
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

ORIGIN  
Query Match 69.1%; Score 817; DB 12; Length 1002;  
Best Local Similarity 98.3%; Fred. No. 3.2e-195;  
Matches 868; Conservative 0; Mismatches 10; Indels 5; Gaps 4;  
Qy 302 CCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCAGCGCTTTGATCACAACCAATTTCCC 361  
Db 1 CCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCAGCGCTTTGATCACAACCAATTTCCC 59  
Qy 362 TCGATGATGTGCTTCTGAGTGTCTGTGAGGAACAATGGGAAGTCTGCCAGCAGAGA 421  
Db 60 TCGATGATGTGCTTCTGAGTGTCTGTGAGGAACAATGGGAAGTCTGCCAGCAGAGA 119  
Qy 422 AAATCTCTGCCAAGCCCAAGCTTGAGTTCCCTGTGTCAGGCGGACCTGTGACCATG 481  
Db 120 AAATCTCTGCCAAGCCCAAGCTTGAGTTCCCTGTGTCAGGCGGACCTGTGACCATG 179  
Qy 482 GAAGCAGAGAGCAAGCCCAAGCTTGAGTTCCCTGTGTCAGGCGGACCTGTGACCATG 541  
Db 180 GAAGCAGAGAGCAAGCCCAAGCTTGAGTTCCCTGTGTCAGGCGGACCTGTGACCATG 239  
Qy 542 GCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATGCTCTGTAGGATGGAGACTGG 601  
Db 240 GCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATGCTCTGTAGGATGGAGACTGG 299  
Qy 602 TGGACGGTGTGCTGTGAAGTCTCAGGAGAGAGTATACATCCCGAGCTCCAGCTGGGC 661  
Db 300 TGGACGGTGTGCTGTGAAGTCTCAGGAGAGAGTATACATCCCGAGCTCCAGCTGGGC 359  
Qy 662 AAAGTCTCCCATGGTGGCTGTATGAGGCGCTGTAGGCGGAGAGAGAGAGAGAGAGTGTG 721  
Db 360 AAAGTCTCCCATGGTGGCTGTATGAGGCGCTGTAGGCGGAGAGAGAGAGAGAGAGTGTG 419  
Qy 722 TTGTTTACCTGGGAACCTTGGAGGGCCCTTCTCTATCCGGGAGAGCCAGACAGAGAGGC 781  
Db 420 TTGTTTACCTGGGAACCTTGGAGGGCCCTTCTCTATCCGGGAGAGCCAGACAGAGAGGC 479  
Qy 782 TCTTACTCTGTGCTGAGTCCGCTCAGCGCGCTGTATCTCTGGGACCGGATCAGACACTAC 841  
Db 480 TCTTACTCTGTGCTGAGTCCGCTCAGCGCGCTGTATCTCTGGGACCGGATCAGACACTAC 539  
Qy 842 AGGATCCACTGCTCTTGACAATGGTGGCTGTATCTCTCAGCGCGCTCACCTTCCCTCA 901

```

Db      540 AGGATCTACTGCTTTGACAAATGGCTGGCTGTATCATCTCACCGGCTCACCTTCCCTCA 599
Qy      902 CTCAGAGCCCTGGTGACCAATTAATCTGAGCTGGCGGATGACATCTGCTCCCTACTCAAG 961
Db      600 CTCAGAGCCCTGGTGACCAATTAATCTGAGCTGGCGGATGACATCTGCTCCCTACTCAAG 659
Qy      962 GAGCCCTGTGCTCTGAGAGGGTGGCCGCTCCCTGGCAAGGATATACCCCTACTCTGTG 1021
Db      660 GAGCCCTGTGCTCTGAGAGGGTGGCCGCTCCCTGGCAAGGATATACCCCTACTCTGTG 719
Qy      1022 ACTGTGAGAGGACACCACTCACTGAAAGAGCTGGACAGCTCCCTCTGTTTCTGAA 1081
Db      720 ACTGTGAGAGGACACCACTCACTGAAAGAGCTGGACAGCTCCCTCTGTTTCTGAA 779
Qy      1082 GCTGCCACA-GGGAGGAGTCTTCTCAGTGAGG--TCTCGGAGTCCCTCAGCTTC 1138
Db      780 ACTGCCACAGGGGAGGAGTCTTCTCAGTGAGGAGTCTCGGGGAGTCCCTCAGCTTC 839
Qy      1139 TACATCAGCCTGAATGACGA-GGCTGTCTCTTTGGATGATGCC 1180
Db      840 TACATCAGCCTGATGACGAGGCTGTCTCTTTGGATGATGCC 882

RESULT 4
BQ052468      1069 bp      mRNA      linear      EST 29-MAR-2002
LOCUS      AGENCOURT 6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
5', mRNA sequence.
ACCESSION      BQ052468
VERSION      BQ052468.1 GI:19811808
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 1069)
JOURNAL      NIH-MGC http://mgi.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2118 row: n column: 13
High quality sequence stop: 681.
FEATURES      Location/Qualifiers
source      1..1069
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
clone="IMAGE:5933772"
tissue_type="natural killer cells, cell line"
lab_host="DH10B (phage-resistant)"
clone_lib="NIH_MGC_106"
note="Organ: blood; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

Query Match      61.8%; Score 731.4; DB 12; Length 1069;
Best Local Similarity 94.0%; Pred No. 1.4e-173;
Matches 824; Conservative 0; Mismatches 1; Indels 52; Gaps 4;

```

```

Qy      310 TGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACAATTTCCCTCGATGAT 369
Db      1   TGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACAATTTCCCTCGATGAT 60
Qy      370 GTGCTTCTGAGTGTCTGTCTGAGGACAATGGGAAGTCTCCCGAGAGAGAAAATCTCT 429
Db      61   GTGCTTCTGAGTGTCTGTCTGAGGACAATGGGAAGTCTGTCCCGAGAGAAAATCTCT 120
Qy      430 GCCAAGCCCAAGCTTGAGTTCTCTGTCCAAAGCCAGGAGACCTGTGACCATGAAGCAGA 489
Db      121   GCCAAGCCCAAGCTTGAGTTCTCTGTCCAAAGCCAGGAGACCTGTGACCATGAAGCAGA 180
Qy      490 GAGAAGCAAGCCACACGCGCTGGCGCTGGCAGTTTCCCGGAGGTGGCCCGCCGAGCT 549
Db      181   GAGAAGCAAGCCACACGCGCTGGCGCTGGCAGTTTCCCGGAGGTGGCCCGCCGAGCT 240
Qy      550 GTCTCTGAGACTCTGGGAGCCATTGACCATCTGCTCTGAGGATGGAGACTGTGTGGACGGT 609
Db      241   GTCTCTGAGACTCTGGGAGCCATTGACCATCTGCTCTGAGGATGGAGACTGTGTGGACGGT 300
Qy      610 GCTGTCTGAAGTCTCAGCAGAGAGATATAATCCAGCGTCCAGTGGGCAAGTCTC 669
Db      301   GCTGTCTGAAGTCTCAGCAGAGAGATATAATCCAGCGTCCAGTGGGCAAGTCTC 360
Qy      670 CCATGGGTGGCTGTATGAGGGCTCTGAGCAGGAGAAAGCAGAGAACTGTGTGTGTACC 729
Db      361   CCATGGGTGGCTGTATGAGGGCTCTGAGCAGGAGAAAGCAGAGAACTGTGTGTGTACC 420
Qy      730 TGGGAACCCCTGGAGGGGCTTCTCATCCGGGAGAGCCAGCAGGAGAGGCTCTTACTC 789
Db      421   TGGGAACCCCTGGAGGGGCTTCTCATCCGGGAGAGCCAGCAGGAGAGGCTCTTACTC 480
Qy      790 TCTGTGAGTCCGCTCAGCCGCTCTGATCTGGGAGTGGGAGTGGATCAGACACTACAGGATCCA 849
Db      481   TCTGTGAGTCCGCTCAGCCGCTCTGATCTGGGAGTGGGAGTGGATCAGACACTACAGGATCCA 540
Qy      850 CTGCTCTGACATGGCTGGCTGTACATCTACCGCGCTCAGCTTCCCTCAGCTCCAGGC 909
Db      541   CTGCTCTGACATGGCTGGCTGTACATCTACCGCGCTCAGCTTCCCTCAGCTCCAGGC 600
Qy      910 CCT-GGTGGAGCAATTAATCTGAGCTGGGAGTGGGAGTGGATGACATCTGTGCTACTCAAGGAGCCCT 968
Db      601   CCTGGTGGAGCAATTAATCTGAG----- 623
Qy      969 GTGCTCTGAGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGC 1028
Db      624   -----GGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGC 671
Qy      1029 AGAGGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCCTGTTTCTGAAGCTGCCA 1088
Db      672   AGAGGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCCTGTTTCTGAAGCTGCCA 731
Qy      1089 CA-GGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGG-AGTCCCTCAGCTCTCATCAG 1146
Db      732   CAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGCTCTCATCAG 791
Qy      1147 CCTGAATGACAGAGCTGTCTCTTTGGATGATGCCCTAG 1183
Db      792   CCTGAATGACAGAGCTGTCTCTTTGGATGATGCCCTAG 828

RESULT 5
BQ178487
LOCUS      60328305F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429896 5',
DEFINITION      mRNA sequence.
ACCESSION      BQ178487
VERSION      BQ178487.1 GI:12685190
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE  
 AUTHORS NIH-MGC http://img.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Email: cgapbs-remail.nih.gov  
 Tissue Procurement: DCTB/DTP  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM10182 row: i column: 01  
 High quality sequence stop: 657.

FEATURES  
 source  
 Location/Qualifiers  
 1..778  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4429896"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_91"  
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 56.9%; Score 673.6; DB 12; Length 778;  
 Best Local Similarity 96.8%; Pred. No. 4.8e-159;  
 Matches 730; Conservative 0; Mismatches 19; Indels 5; Gaps 4;

QY 36 GTGACAGCTCAAGGGCCCTGGCCCTTCCTCCCTGGCTCGGCTGCTGGAGGGT 95  
 Db |||||  
 QY 1 GTGACAGCTCAAGGGCCCTGGG-CTTCCCTCCCTGGCTCGGCTGCTGGAGGGT 59  
 Db |||||  
 QY 96 TCCCGAGTCCGAATCCCTTAAGGAGCATGGGCGAGCTGATCCATCCCTGGGTACAAACT 155  
 Db |||||  
 QY 60 TCCCGAGTCCGAATCCCTTAAGGAGATGGGCGAGCTGATCCATCCCTGGGTACAAACT 119  
 Db |||||  
 QY 156 GTGACTGACAGACAGATGCTAGCTACCCAAACCAACCTAGCTCTCCCTGAGATCC 215  
 Db |||||  
 QY 120 GTGACTGACAGACAGATGCTAGCTACCCAAACCAACCTAGCTCTCCCTGAGATCC 179  
 Db |||||  
 QY 216 TCCCGAGTCCGAATCCCTTAAGGAGCATGGGCGAGCTGATCCATCCCTGGGTACAAACT 275  
 Db |||||  
 QY 180 TCCCGAGTCCGAATCCCTTAAGGAGATGGGCGAGCTGATCCATCCCTGGGTACAAACT 239  
 Db |||||  
 QY 276 GGGCCCCAAGCCCTAAGCTGTCAGCCAGACATGGCTCTCAGCAGAGCTGTCTTCCCA 335  
 Db |||||  
 QY 240 GGGCCCCAAGCCCTAAGCTGTCAGCCAGACATGGCTCTCAGCAGAGCTGTCTTCCCA 299  
 Db |||||  
 QY 336 AGCTTTGATGACAAACCAATTTCCCTCGATGATGCTTCTGAGTGTCTGCTGAGGAA 395  
 Db |||||  
 QY 300 AGCTTTGATGACAAACCAATTTCCCTCGATGATGCTTCTGAGTGTCTGCTGAGGAA 359  
 Db |||||  
 QY 396 CAATGGGAAGTGTGCCAGCAGAGAAATCTCTCCAGCCCAAGCTTGAGTTCTCTCTG 455  
 Db |||||  
 QY 360 CAATGGGAAGTGTGCCAGCAGAGAAATCTCTCCAGCCCAAGCTTGAGTTCTCTCTG 419  
 Db |||||  
 QY 456 TCCAGGCCAGGAGCTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 515  
 Db |||||  
 QY 420 TCCAGGCCA-GGACCTGTGACATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478  
 Db |||||  
 QY 516 TGGGCAAGTTTCCCGGAGGTGCCCGGCGAGCTGTCTGCTGAGACTCGGGGAGCCATTGA 575  
 Db |||||  
 QY 479 TGGGCAAGTTTCCCGGAGGTGCCCGGCGAGCTGTCTGCTGAGACTCGGGGAGCCATTGA 538  
 Db |||||  
 QY 576 CCATCGTCTCTGAGGATGAGACTGGTGGACCGGTGCTGCTGAAAGTCTCAGCAGAGAGT 635  
 Db |||||

Db 539 CCATCGTCTCTGAGGATGGAGACTGGTGGCGGTGCTGTGAAGTCTCAGGAGAGT 598  
 QY 636 ATAACATCCCCAGCGTCCACGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGA 695  
 Db 599 ATAACATCCCCAGCGTCCACGTGGGCAAGTCTCACAATGGTGGCTGTATGAGGGCTGA 658  
 QY 696 GCAGGGAGA-AAGCAGAGGAAGTCTGTGTCTTACCTGGGAACCTCGAGGGGCTTCC-- 752  
 Db 659 GCAGGGAGACACAGCAGAGGAAGTCTGTGTCTTACCTGGGAACCTCGAGGGGCTTCCCT 718  
 QY 753 TCATCCGGGAGAGCCAGACAGGAGGCTCTTA 786  
 Db 719 CATCCCGAAGAGCAGACAGGAGGCTCTTA 752

RESULT 6  
 BQ054265  
 LOCUS  
 DEFINITION AGENCOURT 6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
 5', mRNA sequence.  
 ACCESSION BQ054265  
 VERSION BQ054265.1 GI:19813605  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 986)  
 NIH-MGC http://img.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM2125 row: i column: 12  
 High quality sequence stop: 515.

FEATURES  
 source  
 Location/Qualifiers  
 1..986  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5936339"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and SuperScript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 56.0%; Score 662.6; DB 12; Length 986;  
 Best Local Similarity 98.6%; Pred. No. 3.4e-156;  
 Matches 721; Conservative 0; Mismatches 4; Indels 6; Gaps 5;

QY 120 GATGGGCGAGCTGATCCATCCCTGGTGTACAACTGCTGACAGAGATGCTGAGC 179  
 Db |||||  
 QY 1 GCATGGGCGAGCTGATCCATCCCTGGTGTACAACTGCTGACAGAGATGCTGAGC 60  
 Db |||||  
 QY 180 TACCCAAACCAACCACTAGCTCTCCCTGAAAGATCTCCAGGCTCAGAGAGTCTGGGT 239  
 Db |||||  
 QY 61 TACCCAAACCAACCACTAGCTCTCCCTGAAAGATCTCCAGGCTCAGAGAGTCTGGGT 120  
 Db |||||

QY	240	GTCTAGAGCAAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTTAACCTGTGC	299
Db	121	GTCTAGAGCAAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTTAACCTGTGC	180
QY	300	AGCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTTC	359
Db	181	AGCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTTC	240
QY	360	CCTCGATGATGTGTTCTTGATGCTCTGCTGAGGAAACATGGAAGTCTGCCAGAGAA	419
Db	241	CCTCGATGATGTGTTCTTGATGCTCTGCTGAGGAAACATGGAAGTCTGCCAGAGAA	300
QY	420	GAAATCTTCCCAAGCCCAAGCTTGAGTTCCCTCTGTCCAAGCCAGGACCTGTGACCA	479
Db	301	GAAATCTTCCCAAGCCCAAGCTTGAGTTCCCTCTGTCCAAGCCAGGACCTGTGACCA	360
QY	480	TGGAAGCAGAGAGAAAGCAAGCCACAGCCGTGGCCCTGGGCACTTTCCCGCAGGTGGCC	539
Db	361	TGGAAGCAGAGAGAAAGCAAGCCACAGCCGTGGCCCTGGGCACTTTCCCGCAGGTGGCC	420
QY	540	CGGCCGAGCTGTGCTCAGACTCGGGAGCCATTGACCATCGTCTTGAGGATGGAGCT	599
Db	421	CGGCCGAGCTGTGCTCAGACTCGGGAGCCATTGACCATCGTCTTGAGGATGGAGCT	480
QY	600	GGTGGACGGTCTGCTCTGAAAGTCTCAGGACAGAGATATAACATCCCGACGCTCCACGTG	659
Db	481	GGTGGACGGTCTGCTCTGAAAGTCTCAGGACAGAGATATAACATCCCGACGCTCCACGTG	540
QY	660	GCAAGTCTCCATGGGTGGCTGTATCAGAGGCTGAGCAGGAGAAAGCAGAGGAACTGC	719
Db	541	CCAAGTCTCCATGGGTGGCTGTATCAGAGGCTGAGCAGGAGAAAGCAGAGGAACTGC	600
QY	720	TGTTGTTTACCTGGGAACCCCTGGAGGGG- CCTTCTCTCATCCGGAG- AGCCAGACCAAGGAG	777
Db	601	TGTTGTTTACCTGGGAACCCCTGGAGGGGCTTCTCTCATCCGGAGAAAGCAGACAGGAG	660
QY	778	AGGCTCTTACTCTCTGTCAGT- CGGCTCAGCG- CCTGTCATCCTGGG- ACCGATCA	833
Db	661	AGGCTCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGATCCTGCGGACCGGATCA	720
QY	834	GACACTACAGG	844
Db	721	GACCTACAGG	731
RESULT 7			
BQ053486			
LOCUS			
DEFINITION BQ053486 6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253			
5', mRNA sequence.			
ACCESSION BQ053486.1 GI:19812826			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM Homo sapiens (human)			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
<p>NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a>.  National Institutes of Health, Mammalian Gene Collection (MGC)  Unpublished (1999)  Contact: Robert Strausberg, Ph.D.  Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>  Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  cDNA Library Preparation: Rubin Laboratory  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  DNA Sequencing by: Agencourt Bioscience Corporation  Clone distribution: MGC clone distribution information can be  found through the I.M.A.G.E. Consortium/LLNL at:  <a href="http://image.llnl.gov">http://image.llnl.gov</a>  Plate: LLCM2122 row: 1 column: 06  High quality sequence stop: 394.</p>			

FEATURES		Location/Qualifiers	
source		1..878	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5935253"	
		/tissue_type="natural killer cells, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_106"	
		/note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN		Query Match 49.8%; Score 589; DB 12; Length 878;	
		Best Local Similarity 95.4%; Pred. No. 1.2e-137;	
		Matches 660; Conservative 0; Mismatches 26; Indels 6; Gaps 5;	
QY	125	GGGAGCTGATCCATCCCTGCTGATACAACTGCTGACAGAGATGCTGACTACCC	184
Db	1	GGGAGCTGATCCATCCCTGCTGTA-AAACTGCTGACTGCAGACAGATGCTGATC	59
QY	185	AAACCAACACCTAGCTCTCCCTCAAGATCTCCAGGCTGAGAGAGTCTGGGTGCTCT	244
Db	60	AAACCAACACCTAGCTCTCCCTCAAGATCTCCAGGCTGAGAGAGTCTGGGTGCTCT	119
QY	245	AGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTAACCTGTCAGGCA	304
Db	120	AGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTAACCTGTCAGGCA	179
QY	305	GAGCATGCTCTCAGCAGAGCTGCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCG	364
Db	180	GAGCATGCTCTCAGCAGAGCTGCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCG	239
QY	365	ATGATGCTCTTCTGAGTCTCTGAGGAAACAAATGGAAAGTCTGCCAGCAGAGAAA	424
Db	240	ATGATGCTCTTCTGAGTCTCTGAGGAAACAAATGGAAAGTCTGCCAGCAGAGAAA	299
QY	425	TCTTCCCAAGCCCAAGCTTGAGTTCCTCTGTCAAGCCAGGACCTGTGACCATGAA	484
Db	300	TCTTCCCAAGCCCAAGCTTGAGTTCCTCTGTCAAGCCAGGACCTGTGACCATGAA	359
QY	485	GCAGAGAAAGCAAGGCCACAGCCGTGGCCCTGGCGAGTTTCCCGCAGGTGGCCCGCC	544
Db	360	GCAGAGAAAGCAAGGCCACAGCCGTGGCCCTGGCGAGTTTCCCGCAGGTGGCCCGCC	419
QY	545	GAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGCTGGTG	604
Db	420	GAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGCTGGTG	479
QY	605	ACGCTGCTGCTGAAAGTCTCAGGCAGAGATATTAACATCCCGCAGCTCCAGTGGCAAA	664
Db	480	ACGCTGCTGCTGAAATCTCAGGCAGAGATATTAACATCCCGCAGCTCCAGTGGCAAA	539
QY	665	GTCTCCCATGGTGGCTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGACTCTGTTG	724
Db	540	GTCTCCCATGGTGGCTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGACTCTGTTG	597
QY	725	TTACTCTGGAAACCTTGGAGGGCCCTTCTCAT-CCGGGAGAGCCAGACAGGAGGCT-	782
Db	598	TTACTCTGGAAACCTTGGAGGGCCCTTCTCAT-CCGGGAGAGCCAGACAGGAGGCT-	657
QY	783	CTTACTCTCTGT-CAGTCCGCTCAGCGGCC	813
Db	658	CTAATCTCTGTGTAGTCCGCTTAGCGGCC	689
RESULT 8			

Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
Trust/MRC Building Addenbrookes Hospital Cambridge) whose  
assistance we gratefully acknowledge.  
Please visit our web site for further details.  
URL: <http://genome.sgc.riken.go.jp/>.  
URL: <http://fantom.sgc.riken.go.jp/>.

Location/Qualifiers  
1. .2637  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="FANTOM\_DB:E430023D24"  
/db\_xref="MGI:2427569"  
/db\_xref="taxon:10090"  
/clone="E430023D24"  
/cell\_type="thymic cells"  
/tissue\_type="thymus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="2 days neonate"  
358. .1137  
/note="unnamed protein product; MODULATOR OF ANTIGEN  
RECEPTOR SIGNALING MARS (SPTR)[AAL38196, evidence: FASTV,  
1008ID, 100%length, match=777]"  
Putative"  
/codon\_start=1  
/protein\_id="BAC40495.1"  
/db\_xref="GI:26353730"  
/translations="MGSLSRGKTSPSSSGPQDPVSMOPERHKVTAVALGSFPP  
GEOLRLRGLPEPLTIIREDQWMTVQSVSGREVHMPVSIVAKVAHGLWGLSRE  
AEELLNPGNGAFILISQDTERGCVSLVLSRPASWDRIRHVRIORLONGLWLI  
PLRTFPSLHALVHYVELDAGICCPRLPCVTKLQPLPKGKTPTPPVTVPVSSLNKWK  
LDRLFLFEPASGEASLSLSEGLRESLSYSIISLAEDPLDDA"

ORIGIN

Query Match	47.5%;	Score	561.4;	DB 11;	Length	2637;
Best Local Similarity	78.2%;	Pred. No.	2.4e-130;			
Matches	714;	Conservative	0;	Mismatches	191;	Indels
					8;	Gaps
QY	274	AGGGCCCCAAGCCCTAACTCTCCAGCCAGAGCATGGTCTCAGCAGAGCTGCTTCC	333			
Db						
QY	236	AGGACTTGCAAGGCCCTGACCTGTGGGTGAGTG--TGCACATGGCTGACTACCCCTCAT	293			
Db						
QY	334	CAAGCCTTGTGATGACAAACCAATTTCCCTCGATGATGTGCTTCTGAGTGCTCTGTGTGGG	393			
Db						
QY	294	CAAAACGCTGATGCAAAACCTTCCCTTTCCAGGTTTCAGTGTGTTGTGAGCGTCTGCTG	353			
Db						
QY	394	AACATGGGAGTCTGCCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTC	453			
Db						
QY	354	AGTGATGGAAAGTTTGTCCACGACAGAGGAAAAAACCCT---CCAGCCCCAGCCCCAGCTCTTC	410			
Db						
QY	454	TGTCGAAGCCAGGACCTCTGTACCATGGAAGCAGAGAGAACAAGGCCACAGCCGTGGC	513			
Db						
QY	411	TGGTCCAGACAGAGAACCCGTGTCCATGCAACAGAAAGACACAAGETCACAGCTGTGGC	470			
Db						
QY	514	CCTGGCAGTTTCCCGCAGGCTGGCCGCGCAGCTGTGCTGAGACTCGGGAGCCATT	573			
Db						
QY	471	CCTGGCAGTTTCCACAGAGTGAACAGGCCAGACTATCTCTGAGACTCGGGAGCCCGCT	530			
Db						
QY	574	GACCATCGTCTTGAGGATGGACATGGTGGACGGTGTGTCTGAAGTCTCAGGCAGAGA	633			
Db						
QY	531	GACCATCATCTCTGAGATGGATTTGGTGGACAGTCCAGTCCGAAGTCTCAGGCAGAGA	590			
Db						
QY	634	GTATAACATCCAGCGTCCACGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCCT	693			
Db						
QY	591	GTACCAATCCACAGTGTGTATGTGCTAAAGTCGCCACGGGTGGCTGTACAGGGGCT	650			
Db						
QY	694	GAGCAGGAGAAACGACAGAACTGCTGTGTTTACTCGGAAACCTGGAGGGCCTTCT	753			
Db						
QY	651	GAGCCGGAGAAACCGAGAACTACTCTGTTTACTCGGAAACCCCGAGGGCCTTCT	710			
Db						



prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
2049374  
11042159

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5

## TITLE

The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 2974)

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Ahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

# TITLE JOURNAL

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. 2974  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:5830437K10"  
/db\_xref="MGI:2394046"  
/db\_xref="taxon:10090"  
/clone="5830437K10"  
/sex="male"  
/tissue\_type="thymus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
232. .1011  
/note="unnamed protein product; MODULATOR OF ANTIGEN  
RECEPTOR SIGNALING MARS (SPTR)AAL38196, evidence: PASTY,

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. 2974  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:5830437K10"  
/db\_xref="MGI:2394046"  
/db\_xref="taxon:10090"  
/clone="5830437K10"  
/sex="male"  
/tissue\_type="thymus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
232. .1011  
/note="unnamed protein product; MODULATOR OF ANTIGEN  
RECEPTOR SIGNALING MARS (SPTR)AAL38196, evidence: PASTY,

## FEATURES source

100%ID, 100%length, match=777)  
putative"  
/codon\_start=1  
/protein\_id="BAC27168.1"  
/db\_xref="GI:26326949"  
/translation="MGSLSSRGKTSPPSSSPSSGPDQEPVSMOPERHKVHTAVALGSPPA  
GEARLSRLGRLPLTIIISDGDWMTVQSVSGREVHMPVYVAKVAHGLVGLSSEK  
ARELLLPNGPGAFILRSQTRRGYSLSVRLSPASDRIRHRYIQRLDNGWLYIT  
PRUTFFSLHALVEHSELADGICCPUREPCVQKUGLPKGTDPPTVPTVTSLSNWK  
LDRSLFLFEPASGEASLLSEGLRSLSSYIISLAEDPLDDA"  
ORIGIN  
Query Match 45.5%; Score 538.6; DB 11; Length 2974;  
Best Local Similarity 80.9%; Pred. No. 1.5e-124;  
Matches 653; Conservative 0; Mismatches 148; Indels 6; Gaps 2;  
QY 380 GTGCTCTGCTGAGAAACAATGGGAAGTCTGCCAGCAGAGAAGAAATCTCTGCCAACGCCA 439  
DB 214 GTGAGCGTCTGCTGAGTGATGGGAAGTTTGTCCAGCAGAGGGAAC---CTCCAGCCCC 270  
QY 440 AGTTCAGTTCTCTCTCAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAACAG 499  
DB 271 AGCCCCAGTCTCTCTGTCAGACAGAACCCGTGTCCATGCCAACCAAGAACAG 330  
QY 500 GCCACAGCGTGGCCCTGGGCGAGTTTCCCGCAGAGTGGCCCGCCGAGCTGTCTGAGA 559  
DB 331 GTCACAGCTGTGGCCCTGGGCGAGTTTCCCGCAGAGTGAACAGGCCAGACTATCTCTGAGA 390  
QY 560 CTCGGGGGCCATGACCATCTCTCTGAGGATGAGAGTGGTGGACGGTGTCTGTGA 619  
DB 391 CTCGGGGGCCGCTGACCATCTCTCTGAGGATGAGAGTGGTGGACGGTGTCTGTGA 450  
QY 620 GTCTCAGGAGAGAGTATTAACATCCCGACGCTCCAGCTGGGCAAGTCTCCCATGGGTGG 679  
DB 451 GTCTCAGGAGAGAGTATTAACATCCCGACGCTGGTGTGTGTCTGCTGAGTGGTGG 510  
QY 680 CTGTATGAGGGCTGAGCAGGAGAGAGAGAGAGAGAGTCTCTGTTTACTCTGGGAACCT 739  
DB 511 CTGTATGAGGGCTGAGCAGGAGAGAGAGAGAGAGAGTCTCTGTTTACTCTGGGAACCT 570  
QY 740 GGAGGGGCTCTCTCTCATCTCGGAGAGAGAGAGAGAGAGTCTCTCTCTCTCTGTCAGTC 799  
DB 571 GGAGGGGCTCTCTCTCATCTCGGAGAGAGAGAGAGAGAGTCTCTCTCTCTCTGTCAGTC 630  
QY 800 CGCTCAGCGGCTCTGATCTCTCGGAGAGAGAGAGAGAGAGTCTCTCTCTCTCTGTCAGTC 859  
DB 631 CGACTCAGCGGCTCTGATCTCTCGGAGAGAGAGAGAGAGTCTCTCTCTCTCTGTCAGTC 690  
QY 860 AATGGCTGCTGTACATCTCACCGGCTCTACCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 919  
DB 691 AATGGCTGCTGTACATCTCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 750  
QY 920 CATTAATCTGAGCTGGCGGATGACATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 979  
DB 751 CATTAATCTGAGCTAGCAGATGGCATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 810  
QY 980 AGGGCTGGCGGCTCTCTCTGAGAGAGATATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1039  
DB 811 AGCTTGGGCCACTACTCTGGGAGAGATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 870  
QY 1040 CTCAACTGAAAGAGCTGGACAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1096  
DB 871 CTAAATGGAAGAGAGCTGGACCGCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 930  
QY 1097 GAGTCT 1156  
DB 931 GATCT 990  
1157 GAGGCTGTCT 1183  
DB 991 GACCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017

## CDS



RESULT 11  
 AK020837  
 LOCUS  
 DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930009E21 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING  
 MARS, full insert sequence.  
 AK020837  
 ACCESSION AK020837.1 GI:12861542  
 VERSION  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 JOURNAL 20499374  
 MEDLINE 11042159  
 PUBMED  
 3  
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 JOURNAL 20530913  
 MEDLINE  
 PUBMED 11076861  
 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 685-690 (2001)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 JOURNAL  
 MEDLINE  
 PUBMED 6 (bases 1 to 926)  
 Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hangaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imokani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
 Please visit our web site (http://genome.gsc.riken.go.jp/) for

## further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGATCCCAAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through two rounds of normalization to Rot = 20.0 and subtraction to Rot = 458.8. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTCTTAATTAATATCCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. Retina RNA was provided by Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.

Retina RNA was provided by Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.

## FEATURES

Location/Qualifiers  
 1..926  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM.DB:A930009E21"  
 /db\_xref="MG1:1911678"  
 /db\_xref="taxon:10090"  
 /clone="A930009E21"  
 /tissue\_type="retina"  
 /clone\_libs="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stages="adult"  
 128..667  
 /note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTR)AAL38196, evidence: FASTY, putative."  
 100%ID, 100%length, match=777)  
 /protein\_id="BAB32223.1"  
 /db\_xref="GI:12861543"  
 /translation="MPSVYVAKVAHGWLYEGLSREKAEELLLLFNGPFGALFRESQT RRGCVSLVRLSRPASDRIRHYRIQLDNGWLYISPLTFPSLHALVHYSEADGI CCLPRLPCVLKGLPLGKDPPTPTVPTSSLNWKLDRLSLLFLEAPASGEASLSEGL LRESLSYSVLSLAEDPLDDA"  
 907..912  
 /note="putative"  
 926  
 /note="putative"  
 polyA\_signal  
 polyA\_site

## ORIGIN

Query Match 40.3%; Score 476.4; DB 11; Length 926;  
 Best Local Similarity 83.0%; Pred. No. 3.7e-109;  
 Matches 556; Conservative 0; Mismatches 11; Indels 3; Gaps 1;  
 QY 517 GGGCAGTTTCCCGCAGTGGCCGCGGCGAGCTGCTCGTAGACTCGGGGAGCCATTGAC 576  
 Db 4 GGGCAGTTTCCCGCAGTGGAGCAACAGCCAGACTATCTCTGAGACTCGGGGAGCCGCTGAC 63  
 QY 577 CATCGTCTCTCAGATGGAGACTGGTGGACGGTCTCTCTTCTAGCTCTCAGGCGCCTGAG 636  
 Db 64 CATCATCTCTCAGGATGGAGATTGGTGACAGTCCAGTCGGAAGCTCTCAGCAGAGAGTA 123  
 QY 637 TAAATCCCCAGCGTCCACGTCGGGCAAAAGTCTCCATGGGTGGTGTGTATGAGGCGCTGAG 696  
 Db 124 CCACATGCCAGTGTGTATGTGGCTAAAGTCGCCACCGGTGGCTGTACGAGGGCGCTGAG 183  
 QY 697 CAGGAGAAACACAGAGAACTGTGTTGTTACTTGGGAACCTCGAGGGGCGCTTCTCAT 756  
 Db 184 CCGGAGAAACCGAGGAACCTACTCTGTACCTGGGAACCCCGGAGGGGCGCTTCTCAT 243  
 QY 757 CCGGAGAGCCAGCAGGAGAGGCTCTTACTCTCTCTGTCAGTCCGCTCAGCCCGCCTGC 816





Query Match 38.2%; Score 452.2; DB 13; Length 660;  
 Best Local Similarity 83.4%; Pred. No. 3.9e-103;  
 Matches 538; Conservative 0; Mismatches 103; Indels 4; Gaps 2;

QY 517 GGGCAGTTTCCCGCAGTGGCGCGCGAGTGTGCTGAGACTCGGGAGCCATTGAC 576  
 Db 4 GGGCAGTTTCCCGCAGTGGCGCGCGAGTGTGCTGAGACTCGGGAGCCATTGAC 576

QY 577 CATGCTCTGAGATGGAGTGGTGGAGCGGTGCTGTAAGTCTCAGGACAGAGTA 636  
 Db 64 CATCATCTCTGAGATGGAGTGGTGGAGCGGTGCTGTAAGTCTCAGGACAGAGTA 636

QY 637 TACATCCCGAGTCCACCTGGCGAAGCTCCATGGGTGGTGTATGAGGCGCTGAG 696  
 Db 124 CCATGCCCGAGTGTGATGGCTAAGTGGCCACGGGTGGTGTATGAGGCGCTGAG 696

QY 697 CAGGAGAAAGCAGAGGAATGCTGTTGTTTACCTGGGAACCTCGGAGGCGCTTCAT 756  
 Db 184 CCGGAGAAAGCAGAGGAATGCTGTTGTTTACCTGGGAACCTCGGAGGCGCTTCAT 756

QY 757 CCGGAGAGCCAGCAGGAGGCTCTTACTCTCTGCTCAGTCCGCTCAGCGCGCTGC 816  
 Db 244 CCGGAGAGCCAGCAGGAGGCTCTTACTCTCTGCTCAGTCCGCTCAGCGCGCTGC 816

QY 817 ATCTGGGACCGGATCAGACACTACAGATCCACTGCTTGGCAATGGCTGGCTGACAT 876  
 Db 304 ATCTGGGACCGGATCAGACACTACAGATCCACTGCTTGGCAATGGCTGGCTGACAT 876

QY 877 CTCACCGGCTCAGCTCCGCTCAGTCCAGGCGCTGGTGGACATCTCTGAGCTGC 936  
 Db 364 CTCACCTCGGCTCAGCTCCGCTCAGTCCAGGCGCTGGTGGACATCTCTGAGCTGC 936

QY 937 GGATGACATCTGCTGCTCAAGAGCGCTGTCTCTCAGAGGCGCTGGCGCGCTCCC 996  
 Db 424 AGATGGCATCTGCTGCTCAAGAGCGCTGTCTCTCAGAGGCGCTGGCGCGCTCCC 996

QY 997 TGGCAAGATATACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056  
 Db 484 TGGCAAGATATACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056

QY 1057 GGACAGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1113  
 Db 544 GGACAGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1113

QY 1114 GGGTCTCCGGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1158  
 Db 604 GGGGCTCCAGAT-CCTCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 647

RESULT 13  
 LOCUS BG284179 566 bp mRNA linear EST 21-FEB-2001  
 DEFINITION 602408226F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4520382 5',  
 mRNA sequence.

ACCESSION BG284179  
 VERSION BG284179.1 GI:13034866  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 566)  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>  
 Plate: L1A10418 row: c column: 07  
 High quality sequence start: 2  
 High quality sequence stop: 566.

FEATURES  
 source

1..566  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4520382"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_91"  
 /note="Organ: prostate; vector: pCMV-SPORT6; Site: 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dr primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 34.2%; Score 405; DB 12; Length 566;  
 Best Local Similarity 100.0%; Pred. No. 3e-91;  
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 779 GGCTCTTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 838  
 Db 13 GGCTCTTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 838

QY 839 TACAGGATCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898  
 Db 73 TACAGGATCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898

QY 899 TCACCTCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958  
 Db 133 TCACCTCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958

QY 959 AAGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1018  
 Db 193 AAGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1018

QY 1019 GTGACTGTGAGAGGACACCACTCACTGGAAGAGCTGGAAGAGCTGCTGCTTCT 1078  
 Db 253 GTGACTGTGAGAGGACACCACTCACTGGAAGAGCTGGAAGAGCTGCTGCTTCT 1078

QY 1079 GAAGCTGCCACAGGGAGGAGTCTTCTTCTGAGTGGGTCTCCGGAGTCCCTCAGCTTC 1138  
 Db 313 GAAGCTGCCACAGGGAGGAGTCTTCTTCTGAGTGGGTCTCCGGAGTCCCTCAGCTTC 1138

QY 1139 TACATCAGCCTGATGACGAGGCTGCTCTTTGGATGATGCTAG 1183  
 Db 373 TACATCAGCCTGATGACGAGGCTGCTCTTTGGATGATGCTAG 1183

RESULT 14

LOCUS AL844311 597 bp mRNA linear EST 30-JUL-2002  
 DEFINITION AL844311 pool YT lib v SPD Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AL844311  
 VERSION AL844311.1 GI:22019093  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 597)  
 Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and  
 Sheridan, E.  
 Homo sapiens EST sequence  
 Unpublished (2002)  
 Contact: The Sanger Centre  
 The Sanger Centre  
 Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

Email: humquery@sanger.ac.uk  
Sanger Centre name : scdd10818.154136A  
Homo sapiens EST sequence. This sequence was generated as part of  
The Wellcome Trust Sanger Institute program to identify and  
annotate genes in the human genome. Incomplete or unconfirmed genes  
are experimentally analysed using a variety of cDNA library  
resources. This sequence was obtained from a PCR product generated  
from a pool of up to 100,000 cDNA clones derived from  
pool\_YT.lib\_v\_SPD cDNA library. Further information can be found at  
<http://www.sanger.ac.uk/Teams/Team69/>.

## FEATURES

Location/Qualifiers  
1..597  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="20"  
/clone\_lib="pool\_YT.lib\_v\_SPD"

## ORIGIN

Query Match 27.3%; Score 322.4; DB 9; Length 597;  
Best Local Similarity 99.7%; Pred. No. 2.3e-70;  
Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 860 AATGCTGCTGTACATCTCAACGGGCTCACTTCCCTCACTCCAGGCCCTGTGTGAC 919  
Db 13 AATGCTGCTGTACATCTCAACGGCTCACTTCCCTCACTCCAGGCCCTGTGTGAC 72  
QY 920 CATTAATCTGAGCTGGGATGACATCTGTGCTCACTCAAGGAGCCCTGTGTGAC 979  
Db 73 CATTAATCTGAGCTGGGATGACATCTGTGCTCACTCAAGGAGCCCTGTGTGAC 132  
QY 980 AGGGTGGCCCTCCCTGGCAAGATATACCCCTACCTGTGACGTGTCAGAGACACCA 1039  
Db 133 AGGGTGGCCCTCCCTGGCAAGATATACCCCTACCTGTGACGTGTCAGAGACACCA 192  
QY 1040 CTCACCTGGAAGAGTGAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAG 1099  
Db 193 CTCACCTGGAAGAGTGAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAG 252  
QY 1100 TCTCTTCTCAGTGAGGCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAG 1159  
Db 253 TCTCTTCTCAGTGAGGCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAG 312  
QY 1160 GCTGTCTTTTGATGATGCCTAG 1183  
Db 313 GCTGTCTTTTGATGATGCCTAG 336

RESULT 15  
BU944126 960 bp mRNA linear EST 18-OCT-2002  
LOCUS AGENCOURT\_10545003 NIH\_MGC.107 Homo sapiens cDNA clone  
DEFINITION IMAGE:6728350 5', mRNA sequence.  
ACCESSION BU944126.1 GI:24132945  
VERSION BU944126  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 960)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLCM3049 row: m column: 21  
High quality sequence stop: 628.

## FEATURES

Location/Qualifiers  
1..960  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGH:6728350"  
/issue\_type="adenocarcinoma, cell line"  
/lab\_host="DH103 (phage-resistant)"  
/clone\_lib="NIH\_MGC.107"  
/note="Osgan: breast; Vector: pOTB7; Site 1: EcoRI;  
Site 2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 27.1%; Score 321; DB 13; Length 960;  
Best Local Similarity 88.4%; Pred. No. 7.1e-70;  
Matches 381; Conservative 0; Mismatches 0; Indels 50; Gaps 1;  
QY 753 TCATCCGGAGAGCCAGACAGGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCC 812  
Db 1 TCATCCGGAGAGCCAGACAGGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCC 60  
QY 813 CTGCATCTTGGAGCCGATCAGACATCAGATCCACTGCCTTGACATGGCTGCTGT 872  
Db 61 CTGCATCTTGGAGCCGATCAGACATCAGATCCACTGCCTTGACATGGCTGCTGT 120  
QY 873 ACATCTCACCGCCCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTCTTGAGC 932  
Db 121 ACATCTCACCGCCCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTCT 176  
QY 933 TGGCGGATGACATCTGTGCTTACTCAAGAGCCCTGTGTCTGTGAGAGGCTGGCCCGC 992  
Db 177 -----GAGGGCTGGCCCGC 190  
QY 993 TCCTTGGCAAGGATATACCCCTACCTGTGACTGTGACAGAGGACACCACCTCAACTGGAAAG 1052  
Db 191 TCCTTGGCAAGGATATACCCCTACCTGTGACTGTGACAGAGGACACCACCTCAACTGGAAAG 250  
QY 1053 AGCTGACAGCTCCCTCCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTG 1112  
Db 251 AGCTGACAGCTCCCTCCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTG 310  
QY 1113 AGGCTCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACAGGCTGTCTCTTTGG 1172  
Db 311 AGGCTCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACAGGCTGTCTCTTTGG 370  
QY 1173 ATGATGCCTAG 1183  
Db 371 ATGATGCCTAG 381

Search completed: February 20, 2004, 01:41:31  
Job time : 7840.77 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:32:00 ; Search time 191.966 Seconds  
(without alignments)  
3419.919 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccaaagacc.....tctcttgatgatgctag 1183

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	10.4	2015	4	US-09-023-655-1105
2	119.4	10.1	2298	4	US-09-023-655-1158
3	110	9.3	2354	4	US-09-023-655-1080
4	107.4	9.1	2129	4	US-09-016-434-1452
5	101	8.5	675	1	US-08-707-793A-3
6	101	8.5	675	1	US-08-707-792A-3
7	92.6	7.8	2435	4	US-09-023-655-1313
8	91	7.7	2647	4	US-09-220-132-77
9	91	7.7	2647	5	PCT-US93-06251-77
10	90.8	7.7	1611	1	US-07-820-011A-3
11	90.8	7.7	1611	4	US-09-860-473-3
12	90.8	7.7	1611	5	PCT-US93-00445-3
13	85.6	7.6	1626	4	US-08-860-473-10
14	80.4	6.8	1602	1	US-07-820-011A-1
15	80.4	6.8	1602	5	PCT-US93-00445-1
16	80.4	6.8	1759	4	US-09-470-881-2
17	71	6.0	1491	2	US-09-006-675-1
18	71	6.0	1491	3	US-09-228-603A-1
19	68.4	5.8	282	2	US-09-006-675-5
20	68.4	5.8	282	3	US-09-228-603A-5
21	60.4	5.1	4517	4	US-09-470-881-7
22	60.4	5.1	4517	5	PCT-US93-06251-83
23	59.4	5.0	874	4	US-09-023-655-931
24	55.8	4.7	1467	4	US-09-579-182-2
25	55.8	4.7	1548	4	US-09-099-053-1
26	48	4.1	144	5	PCT-US93-06251-13
27	46	3.9	190	5	PCT-US93-06251-14

28 43.2 3.7 498 6 5219739-21 Patent No. 5219739  
29 42.8 3.6 164 1 US-08-306-691B-28 Sequence 28, Appl  
30 42.8 3.6 164 4 US-09-860-473-17 Sequence 17, Appl  
31 42.8 3.6 164 5 PCT-US93-06251-70 Sequence 18, Appl  
32 41.6 3.5 197 5 PCT-US95-10973A-18 Sequence 70, Appl  
33 41.6 3.5 231 4 US-09-244-583-13 Sequence 13, Appl  
34 41.6 3.5 444 4 US-09-392-932-6 Sequence 6, Appl  
35 41.6 3.5 444 4 US-09-574-708A-1 Sequence 1, Appl  
36 41.6 3.5 444 4 US-09-392-931-1 Sequence 1, Appl  
37 41.6 3.5 456 5 PCT-US95-10973A-88 Sequence 88, Appl  
38 41.6 3.5 467 5 PCT-US95-10973A-86 Sequence 86, Appl  
39 41.6 3.5 473 3 US-08-718-904-1 Sequence 1, Appl  
40 41.6 3.5 473 4 US-09-449-249-1 Sequence 1, Appl  
41 41.6 3.5 473 5 PCT-US95-10973A-25 Sequence 25, Appl  
42 41.6 3.5 495 4 US-09-244-583-25 Sequence 25, Appl  
43 41.6 3.5 495 4 US-09-037-983C-14 Sequence 14, Appl  
44 41.6 3.5 498 6 5194596-20 Patent No. 5194596  
45 41.6 3.5 516 3 US-08-784-551C-1 Sequence 1, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-023-655-1105  
; Sequence 1105, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1105:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2015 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g183911  
US-09-023-655-1105

Query Match 10.4%; Score 123; DB 4; Length 2015;

	Best Local Similarity	53.9%;	Pred. No. 1.5e-25;	Matches 283;	Conservative 0;	Mismatches 230;	Indels 12;	Gaps 1;
QY	457	CAAAGGCCAGGACCTGTGACCATCTGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCT	516					
Db	303	CAACAGCAACACACAGGAATCAGGAGGACGGCTCTGAGGACATCATCTGTGGTTGCCCT	362					
QY	517	GGGCAGCTTCCCGCAGGTGCCCGGCGAGCTCTGCTGAGACTCTGGGAGACCAATTGAC	576					
Db	363	GTATGATTACGAGGCCATTACCAACAGACCTCAGCTTCAGAAGGGGACCAATGCT	422					
QY	577	CATCGTCTCTGAGGATGGAGACTGGTGGACGGTCTGTCTGAAGTCTCAGCAGAGAGTA	636					
Db	423	GGTCTTAGGAATCCGGGNGTGGTAGGCTGATCCCTGGCCACCCGGAAGAGGG	482					
QY	637	TAAATCCCAAGCTTCAAGTGGCAAGTCTCCCAT-----GGGTGGCTGTA	684					
Db	483	CTACATCCCAAGCAACTATGTGCGCCGCGTGTACTCTCTGGAGACAGAGAGTGGTTTT	542					
QY	685	TGAGGGCTTAGCAGGGAGAGAAGCAGAGGAACTGCTGTGTACTCTGGAAACCTCGAGG	744					
Db	543	CAAGGGATCAGCCGGAAGACGCAGAGCGCAACTGTGCTCCCGCAACATGCTGGG	602					
QY	745	GGCCTTCTCATCCGGAGAGCCACAGGAGAGGCTCTTACTCTCTGTCAAGTCCCGCT	804					
Db	603	CTCCTTCATGATCGGGATAGCGAGACCACTAAAGGAAGTACTCTTTGTCCGTGCGAGA	662					
QY	805	CAGCGGCTTCATCTCTGGGACCGGATCAGACATACAGATCCATGCTTGTGACATGG	864					
Db	663	CTACGACCTTCGGCAGGAGATACCGTGAACATTACAAGATCCGACCTGGCAACCG	722					
QY	865	CTGCTCTGATCTACCGCGGCTTCACTTCCCTCACTCCAGGCCCTGGTGGACCAATTA	924					
Db	723	GGGTTCTACATATCCCCCGAAGCACTTCAGCACTCTCGAGAGTGTGTGACCACTA	782					
QY	925	CTCTGAGCTGGCGATGACATCTCTGCTGCTTACTCAAGAGCCCTG	969					
Db	783	CAAGAAGGGAAACAGCGGGTCTGCCAGAACTCTGGTGGCCCTG	827					

RESULT 2  
US-09-023-655-1158  
; Sequence 1158, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Sellhammer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.

```

; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1158:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2298 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G187268
; US-09-023-655-1158

Query Match 10.1%; Score 119.4; DB 4; Length 2298;
Best Local Similarity 53.3%; Pred. No. 1.8e-24;
Matches 283; Conservative 0; Mismatches 236; Indels 12; Gaps 1;

QY 454 TGTCGAAGGCACGGGACCTGTGACCATGGAAGCAGAGAGAAAGCAAGCCACACGCCGTGGC 513
DB 447 TTTACCTGGACAGAGGTTTCAAACTAAAGATCCAGAGGAACAGAGGACATTCGTGTAGC 506
QY 514 CTTGGGCAGTTTCCCGCAGGTGGCCGCGGAGCTGTGCTCAGACTCGGGAGCCATT 573
DB 507 CTTGTACCCCTATGATGGCATCCACCGGACGACTTGTCTTTCAGAAAGGAGAGAAGAT 566
QY 574 GACCATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGTGAAGTCTTCAGGCAGAGA 633
DB 567 GAAAGTCCTGGAGGAGCATGGAGAATGGTGGAAAGCAAAAGTCCCTTTTAACAAAAAAGA 626
QY 634 GTATACATCCCAGCGTCCACTGGCGCAAGGTCTCCC-----ATGGGTGGCT 681
DB 627 AGGCTTCATCCCAGCAACTATGTGGCCAACTCAACACTTAGAAACAGAGAGTGGTT 686
QY 682 GTATGAGGGCTTCAGCAGGAGGAGAAAGCAGAGGAACGTCTGTGTGTATTACCTGGGAACCTGG 741
DB 687 TTTCAAGGATATAACCAGGAAGACCCAGAAAGCGCAGCTTTTGGCACCAAGAAATAGCGC 746
QY 742 AGGGGCTTCCTCATCCGGGAGACCCAGACCCAGGAGGCTCTTACTCTGTCTGCTCAGCG 801
DB 747 TGGAGCTTTCCTTATTAGAGAAAGTGAACATTTAAAGGAAGCTTCTCTGTCTGTCTGAG 806
QY 802 CCTCAGCGCCCTGTCATCTCTGGACCGGATCAGACTACAGGATCCAATGCTTGCACAA 861
DB 807 AGACTTTGACCCCTGTGATGGTGTATTAAAGCACTACAAAATTAGAAGTCTGGATAA 866
QY 862 TGGTGTGCTGTACATCTCACCAGGCTCACCCTCCCTCACCCTCAGGCCCTGGTGACCA 921
DB 867 TGGGGGCTATTACATCTCTCCAGCAATCACTTTTCCCTGTATCAGCAATGATTAACA 926
QY 922 TTACTGTAGCTGGCGATGACATCTGTGCTCTACTCAAGGAGCCCTGTGT 972
DB 927 TTACCAAGAGCGGCAGATGGCTTGTGCAGAGATTGGAGAGGCTTGTAT 977

RESULT 3
US-09-023-655-1080
; Sequence 1080, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA

```

```

; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1080:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2334 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g182573
;
US-09-023-655-1080

Query Match 9.3%; Score 110; DB 4; Length 2354;
Best Local Similarity 56.2%; Pred. No. 9.2e-22;
Matches 234; Conservative 0; Mismatches 170; Indels 12; Gaps 1;

QY 590 GATGAGACTGGTGACGGTCTGCTCAAGTCTCAGGACAGAGATATACATCCCGCAGC 649
DB 478 GAAGTGACTGGTGGAGGCTCGGTCTCTCAGCTCCGGAATACTGGCTGCATTCCCGC 537

QY 650 GTCCACGTGGGCAAGT-----CTCCATGGGTGGCTGTATGAGGGGCTCGAC 697
DB 538 AACTACGTGGCCCTGTTGACTCAATCAAGCTGAGAGTGGTACTTGGNAAGATTGGG 597

QY 698 AGGGAAGACAGAGGAACCTGTTGTTACTCTGGAAACCTCGAGGGGCTTCTCATC 757
DB 598 AGAAAGGATGCAGAGAGCAGCTGCTTTTCAACAGGCAACCCCGAGGGGGCTTCTCAT 657

QY 758 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCGCCCTGCA 817
DB 658 CGGGAAGCGAGACCAACCAAGTGCCTACTCTCTGTCTCCATCCGGGACTGGATCAGAC 717

QY 818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGTCCTTGAATGCTGGCTGTACATC 877
DB 718 AGAGCGCATATGTGAAGCATTACAAGATCCGAAACTGGACATGGCGGCTACTACATC 777

QY 878 TCACCGCCCTCACTTCCCTCACTCCAGGCTGTGGACCACTACTCTGAGCTGGCG 937
DB 778 ACCACCGGTTCACTTCACTCGTGCAGAGCTGTGTGAGCACTACATGAGAGTGAAT 837

QY 938 GATGACATCTGCTGCTACTCAAGAGCCCTGTGTCTCTGAGAGGGGTGCGCCGCT 993
DB 838 GACGGGCTGTCAACCTGCTCATCGCGCCCTGCACCATCATGAAGCCGAGAGGCT 893

RESULT 4
US-09-016-434-1452
; Sequence 1452, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer

```

```

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1452:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9775207
;
US-09-016-434-1452

```

```

Query Match 9.1%; Score 107.4; DB 4; Length 2129;
Best Local Similarity 53.2%; Pred. No. 4.9e-21;
Matches 259; Conservative 0; Mismatches 216; Indels 12; Gaps 1;

QY 510 TGGCCCTGGGCAAGTTTCCCGCAGGTGGCCCGCGAGTGTCTGAGACTCGGGGAGC 569
DB 259 TCGCTCTGCACAGCTATGAGCCCTCTCAGCAGGAGATCTGGGCTTGAAGAGGGGAAC 318

QY 570 CATTGACCATCTCTCTGAGGATGAGAGCTGGTGAACGGTGTCTGTGAAGTCTCAGGA 629
DB 319 AGCTCGCATCTCTGAGCAGAGCGCGGAGTGGTGAAGGCGCAGTCCCTGACCCAGGGGCC 378

QY 630 GAGAGTATAACATCCCGAGCGTCCACGTGGGCAAA-----GTCTCCCATGGGT 677
DB 379 AGGAAGGCTTCATCCCTTCATTTTGTGGCAAGCGACAGCTCGAGCCGACCCCT 438

QY 678 GCGTGTATGAGGGCTGAGCAGGAGAAAGAGAGAACTGCTGTGTGTACTCTGGAAC 737
DB 439 GGTTCCTCAAGAACCTGAGCGCGCAAGGAGCGGAGCGCAGCTCTCTGCGCCCGGGAACA 498

QY 738 CTGGAGGGGCGCTTCTCTCATCCGGGAGAGCCAGACAGGAGAGGCTTACTCTCTGTCTCAG 797
DB 499 CTCACGGCTCTTCTCTCATCCGGGAGAGGAGAGCACCGCGGGATCGTTTTCACGTCTGG 558

QY 798 TCCGCTCAGCGCCCTCTGATCTCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTG 857
DB 559 TCCGGGACTTCGACCAGAACCCAGGAGAGGTGGTGAACATTACAGATCCGTAATCTGG 618

QY 858 ACAATGGCTGGCTGTACATCTCACCGCCCTCACTTCCCTCCCTCACTCCAGGCGCTGGTGG 917
DB 619 ACAAGGTGGCTTCTACATCTCCCTTCGAATCACTTTTCCCGGCTGCTGATGAATGTTCC 678

```







```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Digiglio, Frank S.
;
; REGISTRATION NUMBER: 31,346
;
; REFERENCE/DOCKET NUMBER: 8586
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 516-742-4343
;
; TELEFAX: 516-742-4366
;
; TELEX: 230 901 SANS UR
;
; INFORMATION FOR SEQ ID NO: 77:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 2647 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: double
;
; TOPOLOGY: linear
;
;
; MOLECULE TYPE: DNA (genomic)
;
; PCT-US93-06251-77

```

Query Match	7.7%	Score 91	DB 5	Length 2647
Best Local Similarity	56.5%	Prod. No. 2.9e-16		
Matches 169	Conservative	0	Mismatches 130	Indels 0
Gaps				
QY	672	ATGGTGCTCTATGAGGGCCCTGAGCAGGAGAGAACGAGGAACCTGCTGTGTGTACCTG	731	
Db	1019	AAGAGTGTACTTTTGGAAACTTGGCCGAAAGATGCTGAGCGACAGCTATTGTCTTTG	1078	
QY	732	GGACCTTGGAGGGCCCTTCTCATCTGGGGAGACGACGAGAGAGGCTCTTACTCTC	791	
Db	1079	GAACCCCAAGAGGTACCTTTCTTATCCGGAGAGTGAACCAACCAAGGTCCTATTAC	1138	
QY	792	TGTCAGTCGGCCCTCAGCGGCCCTGCATCTCTGGACCGGATCAGACACTACAGGATCCACT	851	
Db	1133	TTTCTATCCGTGATTGGATGATGAAAGGAGACCATGTCAACATTATAAAATTCGA	1198	
QY	852	GCCTTGACAATGGCTGCTACATCTCACCGGCTCACCTTCCCTCCTACCTCCAGGCC	911	
Db	1199	AAC TTGACATGTTGGATACATATTACACCGGGCCAGTTTGAACACTTCGACAGC	1258	
QY	912	TGGTGGACCAATTACTCTGAGCTGGCGGATGACATCTGTGCTACTCAAGGAGCCCTGT	970	
Db	1259	TTGTACAACATTACTCAGAGAGCTGCAAGTCTCTGTGCGGCTAGTAGTCCCTGT	1317	

RESULT 10

RESOL 10  
US-07-820-011A-3  
; Sequence 3, Application US/07820011A  
; Patent No. 5336615

GENERAL INFORMATION:  
APPLICANT: Leonard Bell;  
APPLICANT: Madri, Joseph A.  
APPLICANT: Warren, Stephen L.  
APPLICANT: Luthringer, Daniel J.  
TITLE OF INVENTION: Genetically Engineered  
Endothelial Cells Exhibiting Enhanced  
Migration  
TITLE OF INVENTION: Migration  
TITLE OF INVENTION: And Proliferation  
TITLE OF INVENTION: Activator Activity

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb storage
;
; COMPUTER: IBM PC XT
; OPERATING SYSTEM: PC-DOS/MS-DOS 2.10
; SOFTWARE: Displaywrite 3

```

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US-07/820,011a  
 FILING DATE: 19920106  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Klee, Maurice M.  
 REGISTRATION NUMBER: 30,399  
 REFERENCE/DOCKET NUMBER: LB-101  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (203) 255 1400  
 TELEFAX: (203) 254 1101  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1611  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: Double  
 TOPOLOGY: Linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapien  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT: Chromosome 20  
 PUBLICATION INFORMATION:  
 AUTHORS: Anderson, Stephen K.  
 AUTHORS: Gibbs, Carol P.  
 AUTHORS: Tanaka, Akio  
 AUTHORS: Kung, Hsing-Jien  
 AUTHORS: Fujita, Donald J.  
 TITLE: Human Cellular src Gene:  
 TITLE: Nucleotide Sequence and Deri  
 TITLE: Acid Sequence of the Region  
 TITLE: the Carboxy-Terminal Two-Th  
 TITLE: pp60c-src  
 JOURNAL: Molecular and Cellular B  
 VOLUME: 5  
 ISSUE: 5  
 PAGES: 1122-1129  
 DATE: May, 1985  
 PUBLICATION INFORMATION:  
 AUTHORS: Tanaka, Akio  
 AUTHORS: Gibbs, Carol P.  
 AUTHORS: Arthur, Richard R.  
 AUTHORS: Anderson, Stephen K.  
 AUTHORS: Kung, Hsing-Jien  
 AUTHORS: Fujita, Donald J.  
 TITLE: DNA Sequence Encoding the  
 TITLE: Amino-Terminal Region of the  
 TITLE: Protein: Implications of Seq  
 TITLE: Divergence among src-Type Ki  
 TITLE: Oncogenes  
 JOURNAL: Molecular and Cellular B  
 VOLUME: 7  
 ISSUE: 5  
 PAGES: 1978-1983  
 DATE: May, 1987  
 US-07-820-011a-3

	Query Match	7.7%;	Score 90.8;	DB 1;	Length 1611;
	Best Local Similarity	53.4%;	Pred. No. 2.6e-16;		
	Matches .222;	Conservative 0;	Mismatches 182;	Indels 12;	Gaps 1;
QY	556	GAGACTCGGGGAGCCATTACCATCGTCTCTGAGGATGAGACTGCTGGACGGTGTGTC	615		
Db	318	GAAAGGAGCGGCTCCGATTGTCAACAACAGAGGGAGACTGCTGGTGGCCCACTC	377		
QY	616	TGAAGTCTCAGGCAGAGATATACTCCCGAGGTCCACGTGG-----GCAA	663		
Db	378	GCTAGCACAGGACAGACAGGCTATATCCCGCAACTACGTGGCGCCCTCCGACTCCAT	437		
QY	664	AGTCTCCCATCGGTGCTGTATGAGGCGCTGAGCAGGGAGAAAGACAGAGAACTGCTGT	723		

```

RESULT 12
PCT-US93-00445-3
Sequence 3 Application PC/TUS9300445
GENERAL INFORMATION:
APPLICANT: Bell Leonard
APPLICANT: Madri, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Luthringer, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 760 Kb storage  
COMPUTER: DELL 486/50  
OPERATING SYSTEM: DOS 5.0  
SOFTWARE: Displaywrite 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00445  
FILING DATE: 19930105  
CLASSIFICATION:  
Prior Application Data:  
APPLICATION NUMBER: 07/820,011  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-101PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255 1400  
TELEFAX: (203) 254 1101  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1611  
TYPE: NUCLEIC ACID  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: Chromosome 20  
PUBLICATION INFORMATION:  
AUTHORS: Anderson, Stephen K.  
AUTHORS: Gibbs, Carol P.  
AUTHORS: Tanaka, Akio  
AUTHORS: Kung, Hsing-Jien  
AUTHORS: Fujita, Donald J.  
TITLE: Human Cellular src Gene:  
TITLE: Nucleotide Sequence and Derived Amino  
TITLE: Acid Sequence of the Region Coding for  
TITLE: the Carboxy-Terminal Two-Thirds of  
TITLE: pp60c-src  
JOURNAL: Molecular and Cellular Biology  
VOLUME: 5  
ISSUE: 5  
PAGES: 1122-1129  
DATE: May, 1985  
PUBLICATION INFORMATION:  
AUTHORS: Tanaka, Akio  
AUTHORS: Gibbs, Carol P.  
AUTHORS: Arthur, Richard R.  
AUTHORS: Anderson, Stephen K.  
AUTHORS: Kung, Hsing-Jien  
AUTHORS: Fujita, Donald J.

;/ TITLE: DNA Sequence Encoding the  
;/ TITLE: Amino-Terminal Region of the Human c-src  
;/ TITLE: Protein: Implications of Sequence  
;/ TITLE: Divergence among src-Type Kinase  
;/ TITLE: Oncogenes  
;/ JOURNAL: Molecular and Cellular Biology  
;/ VOLUME: 7  
;/ ISSUE: 5  
;/ PAGES: 1978-1983  
;/ DATE: May, 1987  
;/ PCT-US93-00445-3

Query Match 7.7%; Score 90.8; DB 5; Length 1611;  
Best Local Similarity 53.4%; Pred. No. 2.6e-16;  
Matches 222; Conservative 0; Mismatches 182; Indels 12; Gaps 1;

QY 556 GAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTGTC 615  
DB 318 GAAAGCGAGCGGCTCCAGATTGTCAACAACAGAGGAGACTGGTGGCCCACTC 377  
QY 616 TGAAGTCTCAGCAGAGAGTATACATCCCGAGCGTCCAGTGG-----GCAA 663  
DB 378 GCTCAGCAGCAGCAGCAGAGCTATCCCGCAGCACTAGCTGGCGCCCTCCGACTCCAT 437  
QY 664 AGTCTCCCATGGGTGGTGTATGAGGGCTTGAGCAGGAGAGAAAGCAGAGAACTGTGTT 723  
DB 438 CCAGCGTGAGAGTGGTATTTTGGCAAGATCACAGAGCGGAGTCAAGCGGTACTGCT 497  
QY 724 GTTACCTGGGAACCTTGAGGGGCTTCTATCCGGAGAGCGACACAGGAGAGGCTC 783  
DB 498 CAATGCAGAGAACCCGAGAGGAGGACCTTCTCGTCCGAGAAAGTGAGACCACGAAAGTGC 557  
QY 784 TTACTCTCTGTCAGTCCCGCTCAGCGCCCTGATGAGGAGAGAAAGCAGAGAACTGTGTT 843  
DB 558 CTACTGCTCTCAGTGTCTGACTTGGACAGCGCAAGGGCTCAACGTGAAGCACTACAA 617  
QY 844 GATCCACTGCTTGACAAATGCTGGTGTATATCTACCGCGCTCACCTTCCCTCACT 903  
DB 618 GATCCGGAAGTGGACAGCGCGGCTTCTATCATCCCTCCGCAACCCAGTTCACAGCCT 677  
QY 904 CCAGGCGCTGTGACCACTACTCTGAGTGGGGATGACATCTGTCGCTACTCA 959  
DB 678 GCACAGCTGGTGCCTACTACTTCCAAACAGCGCGATGGCTGTGCCACCGCCTCA 733

RESULT 13  
US-09-860-473-10  
; Sequence 10, Application US/09860473  
; Patent No. 6656732  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Andrew T. Watt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-C EXPRESSION  
; FILE REFERENCE: RTS-0222  
; CURRENT APPLICATION NUMBER: US/09/860,473  
; CURRENT FILING DATE: 2001-05-18  
; NUMBER OF SEQ ID NOS: 169  
; SEQ ID NO 10  
; LENGTH: 1626  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1626)  
US-09-860-473-10

Query Match 7.6%; Score 89.6; DB 4; Length 1626;  
Best Local Similarity 56.9%; Pred. No. 5.8e-16;  
Matches 164; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 672 ATGGGTGGCTGTATGAGGGCTGAGCAGGAGAGAAAGCAGAGAACTGCTGTGTACCTG 731  
DB 461 AGGAGTGGTACTTTGGCAAGATCACTAGACGGGAATCAGAGCGGCTGCTCAACGGCG 520

QY 732 GGAACCTCGAGGGGCTTCTCTATCCGGGAGAGCCAGACAGGAGGCTTCTACTCTC 791  
DB 521 AGAACCAGGAGGGGACCTTCTCTGAGGAGAGTGAGACCAAAAAGGTGCTACTGCC 580  
QY 792 TGTCACTCGGCTCAGCGCCCTCGATCCTGGGACCGGATCAGACACTACAGATCCACT 851  
DB 581 TCTCTGATCCGACTTCGACATGCCAAGGGGCTAAATGTGAAACTACAGATCCGA 640  
QY 852 GCCTTGACATGGCTGGGTGTATCTCACCAGGCTCACCTTCCCTCCTCCTCAGGCCC 911  
DB 641 AGCTGACAGCGCGGTTTCTACATCACCCTCCGACCCAGTTCACAGCTTCGACGAGC 700  
QY 912 TGGTGGACCACTACTCTGAGCTGGGGATGACATCTGTCGCTACTCA 959  
DB 701 TCGTGGCTTACTCTCCAAACATGCTGATGGCTGTGTCACCGCCTCA 748

RESULT 14  
US-07-820-011A-1  
; Sequence 1, Application US/07820011A  
; Patent No. 5336615  
; GENERAL INFORMATION:  
; APPLICANT: Bell, Leonard  
; APPLICANT: Madri, Joseph A.  
; APPLICANT: Warren, Stephen L.  
; APPLICANT: Luthringer, Daniel J.  
; TITLE OF INVENTION: Genetically Engineered  
; TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced  
; TITLE OF INVENTION: Migration  
; TITLE OF INVENTION: and Plasminogen Activator Activity  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maurice M. Klee  
; STREET: 1951 Burr Street  
; CITY: Fairfield  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 06430  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb storage  
COMPUTER: IBM PC XT  
OPERATING SYSTEM: PC-DOS/MS-DOS 2.10  
SOFTWARE: Displaywrite 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,011A  
FILING DATE: 19920106  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: LB-101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255 1400  
TELEFAX: (203) 254 1101  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1602 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE:  
ORGANISM: Gallus, gallus  
PUBLICATION INFORMATION:  
AUTHORS: Takeya, Tatsuo  
AUTHORS: Hanafusa, Hidesaburo  
TITLE: Structure and Sequence of the  
TITLE: Cellular Gene Homologous to the RSV src  
TITLE: Gene and the Mechanism for Generating the  
TITLE: Transforming Virus



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:23:34 ; Search time 2843.55 Seconds  
(without alignments)  
1456.787 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctcaagacc.....tctcttgatgagctag 1183

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	784.4	66.3	786	14	US-10-043-649-1
2	775.4	65.5	864	10	US-09-814-353-21302
3	758.2	64.1	763	9	US-09-867-550-953
4	348	28.4	444	9	US-08-867-550-951
5	341	28.8	875	9	US-09-867-550-1915
6	213.4	18.0	320	10	US-09-814-353-17314
7	157.4	13.3	2665	9	US-09-954-456-499
8	157.4	13.3	3756	13	US-10-002-600-91
9	141.8	12.0	432	9	US-09-864-761-2829
10	141.8	12.0	448	9	US-09-864-761-15513
11	131.6	11.1	152	10	US-09-814-353-4631
12	131.6	11.1	152	10	US-09-814-353-10930
13	124.6	10.5	2343	15	US-10-062-674-2038
14	123	10.4	1924	15	US-10-193-720-1
15	123	10.4	2015	9	US-09-954-456-1983

16	123	10.4	2015	14	US-10-007-010-3	Sequence 3, Appl
17	123	10.4	2341	14	US-10-252-157-140	Sequence 140, App
18	119.4	10.1	2298	14	US-10-175-523-50	Sequence 50, Appl
19	119.4	10.1	2298	15	US-10-159-563-343	Sequence 343, App
20	110.6	9.3	2032	15	US-10-366-288-27	Sequence 27, Appl
21	110	9.3	1590	15	US-10-085-117-18	Sequence 18, Appl
22	110	9.3	2354	9	US-09-967-768A-300	Sequence 300, App
23	110	9.3	2354	15	US-10-353-690-123	Sequence 123, App
24	110	9.3	2354	15	US-10-085-117-17	Sequence 17, Appl
25	110	9.3	2433	14	US-10-240-965-114	Sequence 114, App
26	107.4	9.1	2017	15	US-10-062-674-1776	Sequence 1776, App
27	107.4	9.1	2129	10	US-09-960-706-954	Sequence 954, App
28	107.4	9.1	2129	15	US-10-305-720-1452	Sequence 1452, App
29	102.6	8.7	1911	9	US-09-917-800A-1611	Sequence 1611, App
30	99	8.4	1554	15	US-10-088-117-15	Sequence 15, Appl
31	99	8.4	2179	15	US-10-085-117-14	Sequence 14, Appl
32	96	8.1	96	9	US-09-864-761-19612	Sequence 19612, A
33	91	7.7	549	10	US-09-918-995-25044	Sequence 25044, A
34	91	7.7	1609	9	US-09-771-161A-30	Sequence 30, Appl
35	91	7.7	1995	9	US-09-771-161A-31	Sequence 31, Appl
36	91	7.7	4414	14	US-10-101-510-512	Sequence 512, App
37	89.6	7.6	2451	9	US-09-771-161A-4	Sequence 4, Appl
38	84.8	7.2	488	10	US-09-918-995-19745	Sequence 19745, A
39	83.8	7.1	2173	15	US-10-094-749-1087	Sequence 1087, App
40	76	6.4	342	15	US-10-062-674-1242	Sequence 1242, App
41	73.2	6.2	486	9	US-09-796-692-7419	Sequence 7419, App
42	73.2	6.2	486	14	US-10-040-862-7419	Sequence 7419, App
43	73.2	6.2	486	15	US-10-057-475B-7419	Sequence 7419, App
44	73.2	6.2	486	15	US-10-154-884B-7419	Sequence 7419, App
45	60.4	5.1	4343	15	US-10-153-563-184	Sequence 184, App

## ALIGNMENTS

### RESULT 1

US-10-043-649-1  
; Sequence 1, Application US/10043649  
; Publication No. US20030059924A1  
; GENERAL INFORMATION:  
; APPLICANT: Holland, Sacha J.  
; APPLICANT: Mendenhall, Marcy K.  
; APPLICANT: Pardo, Jorge  
; APPLICANT: Spencer Collin  
; APPLICANT: Fu, C. Alan  
; APPLICANT: Luo, Ying  
; APPLICANT: Payan, Donald G.  
; APPLICANT: Mancebo, Helena S.Y.  
; APPLICANT: Wu, Jun  
; APPLICANT: Zhou, Xiulan  
; APPLICANT: Shen, Mary  
; APPLICANT: Liao, X. Charlene  
; APPLICANT: Sheng, Ning  
; TITLE OF INVENTION: Cloning of a No. US20030059924A1  
; FILE REFERENCE: A-70219-1/RMS/DHR  
; CURRENT APPLICATION NUMBER: US/10/043,649  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/260,953  
; PRIOR FILING DATE: 2001-01-10  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 786  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(786)  
; OTHER INFORMATION:  
US-10-043-649-1

Query Match 66.3%; Score 784.4; DB 14; Length 786;

```

Best Local Similarity 99.9%; Pred. No. 1.5e-229;
Matches 795; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGGAAGTGTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTCTCTCTGTC 457
Db 1 ATGGGAAGTGTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTCTCTCTGTC 60
QY 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGACAGAGCCAGCGGTGGCCCTG 517
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGACAGAGCCAGCGGTGGCCCTG 120
QY 518 GCGAGTTTCCCGGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 577
Db 121 GCGAGTTTCCCGGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
QY 578 ATCGTCTCTGAGGATGGAGCTGTGAGCAGTGTCTGAGTCTCAGGACAGAGTAT 637
Db 181 ATCGTCTCTGAGGATGGAGCTGTGAGCAGTGTCTGAGTCTCAGGACAGAGTAT 240
QY 638 AACATCCCAGCGTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGCGCTGAGC 697
Db 241 AACATCCCAGCGTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGCGCTGAGC 300
QY 698 AGGAGAAAGCAGAGAACTCTGTGTTACTGGGAACCTGAGGAGGCTTCTTCATC 757
Db 301 AGGAGAAAGCAGAGAACTCTGTGTTACTGGGAACCTGAGGAGGCTTCTTCATC 360
QY 758 CGGAGAGCCAGACAGAGAGAGGTCTTACCTCTGTCTGACGTCGCCCTCAGCGGCCCTGCA 817
Db 361 CGGAGAGCCAGACAGAGAGAGGTCTTACCTCTGTCTGACGTCGCCCTCAGCGGCCCTGCA 420
QY 818 TCCTGGGACCGATCAGACACTAGAGATCCACTGCGCTTGACAAATGGCTGGCTGTATC 877
Db 421 TCCTGGGACCGATCAGACACTAGAGATCCACTGCGCTTGACAAATGGCTGGCTGTATC 480
QY 878 TCACGCGCCCTCACCTTCCCTCTACTCCAGGCGCTGTGGACCAATTAATCTGAGCTGGCG 937
Db 481 TCACGCGCCCTCACCTTCCCTCTACTCCAGGCGCTGTGGACCAATTAATCTGAGCTGGCG 540
QY 938 GATGACATCTGCTGCTACTCAAGAGCGCTGTGCTCTGAGAGGGGTGCCCGCTCCCT 997
Db 541 GATGACATCTGCTGCTACTCAAGAGCGCTGTGCTCTGAGAGGGGTGCCCGCTCCCT 600
QY 998 GCGAAGATATACCCCTACCTGTGACGTGTGAGAGGACACACTCAACTGGAAGAGCTG 1057
Db 601 GCGAAGATATACCCCTACCTGTGACGTGTGAGAGGACACACTCAACTGGAAGAGCTG 660
QY 1058 GACAGCTCCCTCTGTTTCTGAAAGCTGCCAAGGGAGGAGTCTCTTCTCAGTGAGGGT 1117
Db 661 GACAGCTCCCTCTGTTTCTGAAAGCTGCCAAGGGAGGAGTCTCTTCTCAGTGAGGGT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGACGAGGCTGTCTCTTTGGATGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGACGAGGCTGTCTCTTTGGATGAT 780
QY 1178 GCCTAG 1183
Db 781 GCCTAG 786

```

RESULT 2  
 US-09-814-353-21302  
 ; Sequence 21302, Application US/09814353  
 ; Publication No. US20030165831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thompson, Pamela  
 ; APPLICANT: Lillie, James  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; THERAPY OF OVARIAN CANCER  
 ; FILE REFERENCE: MRI-006B  
 ; CURRENT APPLICATION NUMBER: US/09/814,353

```

CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21302
LENGTH: 864
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 2, 3, 32, 862, 863, 864
OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21302

```

```

Query Match 65.5%; Score 775.4; DB 10; Length 864;
Best Local Similarity 99.2%; Pred. No. 8.7e-227;
Matches 779; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GTGAGGTCCAGGACCCCAAGCCCTGTGTCTGTGTGACAGAGCTCAAAGGGCCCTGGGC 61
Db 54 GTGAGGTCCAGGACCCCAAGCCCTGTGTGTGTGTGACAGAGCTCAAAGGGCCCTGGGC 113
QY 62 CTTCCCTCCCTGGCTGGCTGTGTGTGGAGGTTCCCGAGTCCAGATCCCTTAGGAGC 121
Db 114 CTTCCCTCCCTGGCTGGCTGTGTGGAGGTTCCCGAGTCCAGATCCCTTAGGAGC 173
QY 122 ATGGGGAGCTGATCCATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 181
Db 174 ATGGGGAGCTGATCCATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 233
QY 182 CCAAAACCAACACCTGAGCTCTCCCTGAGATCTCCCGAGTCCAGAGTCTGTGGTGT 241
Db 234 CCAAAACCAACACCTGAGCTCTCCCTGAGATCTCCCGAGTCCAGAGTCTGTGGTGT 293
QY 242 CCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGGCCCTAACCTGTCCAG 301
Db 294 CCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGGCCCTAACCTGTCCAG 353
QY 302 CCAGAGCATGGCTCTCAGCAGAGCTGTCTCCAGAGCCCTTGTGTGTGTGTGTGTGTGT 361
Db 354 CCAGAGCATGGCTCTCAGCAGAGCTGTCTCCAGAGCCCTTGTGTGTGTGTGTGTGTGT 413
QY 362 TCGATGATGTGTCTCTGAGTGTCTCTGAGGAAACAATGGGAAGTCTGCCAGCAGAGA 421
Db 414 TCGATGATGTGTCTCTGAGTGTCTCTGAGGAAACAATGGGAAGTCTGCCAGCAGAGA 473
QY 422 AAATCTCTGCAAGCCCAAGCTTGAATCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 481
Db 474 AAATCTCTGCAAGCCCAAGCTTGAATCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 533
QY 482 GAAGCAGAGAGAAGCAGGCCACAGCGCTGGGCCCTGGGCAAGTCTCCCGGAGGTGGCC 541
Db 534 GAAGCAGAGAGAAGCAGGCCACAGCGCTGGGCCCTGGGCAAGTCTCCCGGAGGTGGCC 593
QY 542 CCGGAGCTGTCTGTGAGACTCGGGAGCCATGACCAATGCTCTGTCTGTCTGTCTGTCTGT 601
Db 594 CCGGAGCTGTCTGTGAGACTCGGGAGCCATGACCAATGCTCTGTCTGTCTGTCTGTCTGT 653
QY 602 TGGAGCGTGTCTGTGAGTCTCAGGAGAGATTAACATCCCGAGCTCCAGCTGGC 661
Db 654 TGGAGCGTGTCTGTGAGTCTCAGGAGAGATTAACATCCCGAGCTCCAGCTGGC 713

```

QY 662 AAGTCTCCATCGGTGGTGGTCTGTATGAGGGCTTGACGAGGGAAGAGAGAGAACTCTG 721  
Db 714 AAGTCTCCATCGGTGGTGGTCTGTATGAGGGCTTGACGAGGGAAGAGAGAGAACTCTG 773  
QY 722 TTGTTACCTGGGAACCTCTGAGGGGCTTCTCTATCCGGAGAGCCAGCCAGAGGC 781  
Db 774 TTGTTACCTGGGAACCTCTGAGGGGCTTCTCTATCCGGAGAGCCAGCCAGAGGC 833  
QY 782 TCTTA 786  
Db 834 TCTTA 838

## RESULT 3

US-09-867-550-953  
; Sequence 953, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; PRIOR FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 953  
; LENGTH: 763  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-867-550-953

Query Match 64.1%; Score 758.2; DB 9; Length 763;  
Best Local Similarity 99.6%; Pred. No. 1.5e-221;  
Matches 760; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 113 CTAAGAGATGGGGAGCTGATGATCCCTGGTGTAACAACCTGCTGACGACAGAT 172  
Db 1 CTAAGAGATGGGGAGCTGATGATCCCTGGTGTAACAACCTGCTGACGACAGAT 60  
QY 173 GCTGAGCTACCAACCAACCACTAGCTCTCCCTGAAGATCCTCCAGGCTGAGAGT 232  
Db 61 GCTGAGCTACCAACCAACCACTAGCTCTCCCTGAAGATCCTCCAGGCTGAGAGT 120  
QY 233 TGTGGTGTCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAA 292  
Db 121 TGTGGTGTCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAA 180  
QY 293 CTTGTCAGCAGAGATCGCTCTCAGCAGAGCTGTCTTCCCAAGCTTTTGTATGACAAAC 352  
Db 181 CTTGTCAGCAGAGATCGCTCTCAGCAGAGCTGTCTTCCCAAGCTTTTGTATGACAAAC 240  
QY 353 CAATTTCCCTCGATGATGCTTCTGAGTGTCTCTGCTGAGGAAACAATGGGAAGTCTGCC 412  
Db 241 CAATTTCCCTCGATGATGCTTCTGAGTGTCTCTGCTGAGGAAACAATGGGAAGTCTGCC 300  
QY 413 AGCAGAAGAAATCTCTGCAAGCCCAAGCTTGTAGTCTCTGCTGCAAGGCCAGGACCT 472  
Db 301 AGCAGAAGAAATCTCTGCAAGCCCAAGCTTGTAGTCTCTGCTGCAAGGCCAGGACCT 360  
QY 473 GTGACCATGAGAGCAGAGAGAGAGAGGACCAAGCCGTGGCCCTGGGAGTTTCCCGCA 532  
Db 361 GTGACCATGAGAGAGAGAGAGAGGACCAAGCCGTGGCCCTGGGAGTTTCCCGCA 420  
QY 533 GGTGGCCCGCCGAGCTGTCTGAGACTCGGGAGCCATTTGACCATGCTCTCTGAGAT 592  
Db 421 GGTGGCCCGCCGAGCTGTCTGAGACTCGGGAGCCATTTGACCATGCTCTCTGAGAT 480

## RESULT 4

US-09-867-550-951  
; Sequence 951, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 951  
; LENGTH: 444  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-867-550-951

Query Match 29.4%; Score 348; DB 9; Length 444;  
Best Local Similarity 100.0%; Pred. No. 4e-96;  
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 CTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTCCAG 301  
Db 1 CTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTCCAG 60  
QY 302 CCAGAGCATGCTCTCAGCAGAGCTGTCTCCCAAGCTTTGTATGACAAACCAATTTCCC 361  
Db 61 CCAGAGCATGCTCTCAGCAGAGCTGTCTCCCAAGCTTTGTATGACAAACCAATTTCCC 120  
QY 362 TCGATGATGTCTTCTGAGTGTCTCTGCTGAGGAAACAATGGGAAGTGTGCCAGAGAAGA 421  
Db 121 TCGATGATGTCTTCTGAGTGTCTCTGCTGAGGAAACAATGGGAAGTGTGCCAGAGAAGA 180  
QY 422 AAATCTCTGCCAAGCCCAAGCTTCTGCTTCCCAAGCTTTGTATGACAAACCAATTTCCC 481  
Db 181 AAATCTCTGCCAAGCCCAAGCTTCTGCTTCCCAAGCTTTGTATGACAAACCAATTTCCC 240  
QY 482 GAAGCAGAGAGAGAGAGAGGACCAAGCCGTGGCCCTGGGAGTTTCCCGAGGTGGCCG 541  
Db 241 GAAGCAGAGAGAGAGAGGACCAAGCCGTGGCCCTGGGAGTTTCCCGAGGTGGCCG 300  
QY 542 GCCGAGCTGTCTGCTGAGACTCGGGAGCCATTGACCATCTCTCTGAG 589

Db 301 GCCGAGCTGCTGCTGAGACTCGGGGAGCCATTGACCATCGTCTCTGAG 348

RESULT 5

US-09-867-550-1915  
 ; Sequence 1915, Application US/09867550  
 ; Patent No. US20020082206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad,  
 ; APPLICANT: Conley, Pamela  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
 ; FILE REFERENCE: 21402-013 (Cura-3113)  
 ; CURRENT APPLICATION NUMBER: US/09/867,550  
 ; PRIOR FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: USN 60/208,427  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1915  
 ; LENGTH: 875  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Wherein n is one of a or t or c or g  
 US-09-867-550-1915

Query Match 28.8%; Score 341; DB 9; Length 875;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-94;  
 Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 843 GGATCCACTGCTGACAAATGGCTGGGTGATCATCTCACCGCGCTCACCTTCCCTCTAC 902  
 Db 2 GGATCCACTGCTGACAAATGGCTGGGTGATCATCTCACCGCGCTCACCTTCCCTCTAC 61  
 Qy 903 TCCAGGCGCTGGTGACATTAATCTGAGTGGGAGTATCATCTGCTGCTACTCAAG 962  
 Db 62 TCCAGGCGCTGGTGACATTAATCTGAGTGGGAGTATCATCTGCTGCTACTCAAG 121  
 Qy 963 AGCCCTGTGCTTCAGAGGCTGGCGGCTCCTGCAAGGATATACCCCTACCTGTCA 1022  
 Db 122 AGCCCTGTGCTTCAGAGGCTGGCGGCTCCTGCAAGGATATACCCCTACCTGTCA 181  
 Qy 1023 CTGTGCAAGGACACCACTCAACTGGAAGAGTGGAGCTCCCTCCTGTTTCTGAAG 1082  
 Db 182 CTGTGCAAGGACACCACTCAACTGGAAGAGTGGAGCTCCCTCCTGTTTCTGAAG 241  
 Qy 1083 CTGCCACAGAGGAGTCTTCTCAGTCAAGGCTCCCGGAGTCCCTCAGCTTCTACA 1142  
 Db 242 CTGCCACAGAGGAGTCTTCTCAGTCAAGGCTCCCGGAGTCCCTCAGCTTCTACA 301  
 Qy 1143 TCAGCTGAATGAGGCTGCTCTTCTTGGATGATGCTTAG 1183  
 Db 302 TCAGCTGAATGAGGCTGCTCTTGGATGATGCTTAG 342

RESULT 6

US-09-814-353-17314  
 ; Sequence 17314, Application US/09814353  
 ; Publication No. US20030165831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lee, John  
 ; APPLICANT: Thompson, Pamela  
 ; APPLICANT: Lillie, James  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
 ; FILE REFERENCE: MRI-006B

; CURRENT APPLICATION NUMBER: US/09/814,353  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/191,031  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/207,124  
 ; PRIOR FILING DATE: 2000-08-25  
 ; PRIOR APPLICATION NUMBER: US 60/211,940  
 ; PRIOR FILING DATE: 2000-06-15  
 ; PRIOR APPLICATION NUMBER: US 60/216,820  
 ; PRIOR FILING DATE: 2000-07-07  
 ; PRIOR APPLICATION NUMBER: US 60/220,661  
 ; PRIOR FILING DATE: 2000-07-25  
 ; PRIOR APPLICATION NUMBER: US 60/257,672  
 ; PRIOR FILING DATE: 2000-12-21  
 ; NUMBER OF SEQ ID NOS: 22037  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 17314  
 ; LENGTH: 320  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-814-353-17314

Query Match 18.0%; Score 213.4; DB 10; Length 320;  
 Best Local Similarity 95.2%; Pred. No. 5.5e-55;  
 Matches 220; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 536 GGCCCGCGCGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGATGGA 595  
 Db 90 GTCCGCGCGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGATGGA 149  
 Qy 596 GACTGTGGACGGTCTCTCTGAAGTCTCAGGAGAGATTAACATCCCGAGGTCCAC 655  
 Db 150 GACTGTGGACGGTCTCTCTGAAGTCTCAGGAGAGATTAACATCCCGAGGTCCAC 209  
 Qy 656 GTGGGCAAGTCTCCCATGGTGGCTGATGAGGCGCTGACGAGGAGAAAGCAGAGAA 715  
 Db 210 GTGGGCAAGTCTCCCATGGTGGCTGATGAGGCGCTGACGAGGAGAAAGCAGAGAA 269  
 Qy 716 CTGCTGTTGTTTACCTGGGAACCTCGAGGGGCTTCTCTCATCCGGAGAGC 766  
 Db 270 CTGCTGTTGTTTACCTGGGAACCTCGAGGGGCTTCTCTCATCCGGAGAGC 320

RESULT 7

US-09-954-456-499  
 ; Sequence 499, Application US/09954456  
 ; Patent No. US20020115057A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
 ; FILE REFERENCE: 689290-76  
 ; CURRENT APPLICATION NUMBER: US/09/954,456  
 ; CURRENT FILING DATE: 2001-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/233,617  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/234,052  
 ; PRIOR FILING DATE: 2000-09-20  
 ; PRIOR APPLICATION NUMBER: US/60/234,923  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235,134  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235,637  
 ; PRIOR FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US/60/235,638  
 ; PRIOR FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US/60/235,711  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,720  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,840  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,863



; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 499  
; LENGTH: 2665  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-499

Query Match 13.3%; Score 157.4; DB 9; Length 2665;  
Best Local Similarity 54.2%; Pred. No. 1.2e-37;  
Matches 354; Conservative 0; Mismatches 281; Indels 18; Gaps 1;  
QY 410 CCACGACAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTCTCAAGCCAGGGA 469  
Db 24 CCAGGAAAAAGAAAGAAATGGGAAACAGCATGAATCCACCCCTGGCGCTCCCGAGAGG 83  
QY 470 CCTGTGACCATGAAGCAGAGAGAACCAAGGCCACAGCCGTCGGCCCTGGGCAAGTTCCCG 529  
Db 84 CCCTGCCCAACCCGGAGGACTGGATAGCGACTTCCTTGGCGTGTCTAAGTGACTACCCG 143  
QY 530 GCAGGTGGCCGCGCGAGCTGTCTGAGACTCGGGAGCCATTGACCATCTCTCTGAG 589  
Db 144 TCTCTTGACATCAGCCCCCGATATTTCCGCCGAGGGGAGAACTGGGTGTGATTCTGAT 203  
QY 590 GATGGAGACTGTGGACGCTGTCTGAGTCTCAGGACAGAGATTAACAATCCCGAGC 649  
Db 204 GAAGGGGCTGTGGAAAGCTATTCTTTAGCACTGGTCGAGAGATTATATCCCTCGA 263  
QY 650 GTCCAGTGGGCAAGTCTCCCATGGGTGGCTGTATGAGGGCTCAGCAGGAGGAGAAACA 709  
Db 264 ATATGTGTGGCCAGAGTTTACCATGGCTGGCTGTTTGGGGCTCGGACAGAGACAGGCC 323  
QY 710 GAGGAACTGCTGTTTACTCTGTCTGTGAGTCCGCTCAGCCGCTGATCTCCGAGAGCCAG 769  
Db 324 GAGGAGCTGCTCAGCTGCGACACAAAGTTCGGCTCTTTCATGATCAGAGAGAGTGAG 383  
QY 770 ACCAGGAGAGGCTTACTCTGTCTGTGAGTCCGCTCAGCCGCTGATCTCCGAGAGCCG 829  
Db 384 ACCAAGAAAGGTTTACTCATCTGCTGAGACAGGCA-----G 425  
QY 830 ATCAGACTACAGGATCAGTCCTGACATGCTGGCTGTATCATCTCACCCGCGCTC 889  
Db 426 GTAAAGCAATTACCGATTTTCGCTGCGCAACAACTGGTACTCATTTTCCCGAGGCTC 485  
QY 890 ACCTTCCCTCACTCAGGCTGTGTGAGCACTTACTCTGAGCTGGCGGATGATCTGC 949  
Db 486 ACCTTCCAGTGCCTGGAGACTGTGTGACCACTATTCTGAGGTGGCTGTGATGGCTGTGC 545  
QY 950 TGCCTACTCAAGAGCCCTGTGTCTGAGAGGGCTGGCCGCTCCCTGGCAAGGATATA 1009  
Db 546 TGTGTGCTCACACGCCCTGCTGACACAAAGCAGCGCTGCCAGCAGTGAGGGCTCC 605  
QY 1010 CCCCTACCTGTGCTGACAGGACACCACTCACTGGAAGAGCTGGACAG 1062  
Db 606 AGCTCACCTGTACCTTGGTCAGAGACTGTGACTGGAGAGAGTGTCCAG 658

## RESULT 8

US-10-002-600-91  
; Sequence 91, Application US/10002600  
; Publication No. US20020137077A1  
; GENERAL INFORMATION:  
; APPLICANT: Hopkins, Christopher M.  
; APPLICANT: Peterson, David P.  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Hawkins, Phillip R.  
; TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS  
; FILE REFERENCE: PA-0042 US  
; CURRENT APPLICATION NUMBER: US/10/002.600  
; PRIOR FILING DATE: 2001-10-25  
; PRIOR APPLICATION NUMBER: 60/243,521  
; PRIOR FILING DATE: 2000-10-25

; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PERL Program  
; SEQ ID NO 91  
; LENGTH: 3756  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Template ID: 059263.15  
US-10-002-600-91

Query Match 13.3%; Score 157.4; DB 13; Length 3756;  
Best Local Similarity 54.2%; Pred. No. 1.4e-37;  
Matches 354; Conservative 0; Mismatches 281; Indels 18; Gaps 1;  
QY 410 CCACGACAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTCTCAAGCCAGGGA 469  
Db 1098 CCAGGAAAAAGAAAGAAATGGGAAACAGCATGAATCCACCCCTGGCGCTCCCGAGAGG 1157  
QY 470 CCTGTGACCATGAAGCAGAGAGAACCAAGGCCACAGCCGTCGGCCCTGGGCAAGTTCCCG 529  
Db 1158 CCCTGCCCAACCCGGAGGACTGGATAGCGACTTCCTTGGCGTGTCTAAGTGACTACCCG 1217  
QY 530 GCAGGTGGCCGCGCGAGCTGTCTGAGACTCGGGAGCCATTGACCATCTCTCTGAG 589  
Db 1218 TCTCTGACATCAGCCCCCGATATTTCCGCCGAGGGGAGAACTGGGTGTGATTCTGAT 1277  
QY 590 GATGGAGACTGTGGACGCTGTCTGAGTCTCAGGACAGAGATTAACAATCCCGAGC 649  
Db 1278 GAAGGGGCTGTGGAAAGCTATTCTTTAGCACTGGTCGAGAGAGTTATCATCCCTCGA 1337  
QY 650 GTCCAGTGGGCAAGTCTCCCATGGGTGGCTGTATGAGGGCTCAGCAGGAGGAGAAACA 709  
Db 1338 ATATGTGTGGCCAGAGTTTACCATGGCTGGCTGTTTGGGGCTCGGACAGACAGGCC 1397  
QY 710 GAGGAACTGCTGTGTGAGTCTGAGGAACTCTGAGAGGGCTTCTCATCCGGAGAGCCAG 769  
Db 1398 GAGGAGCTGCTGAGCTGCGCAGACAAAGTTCGGCTCTTTCATGATCAGAGAGAGTGAG 1457  
QY 770 ACCAGGAGAGGCTTACTCTGTCTGAGTCCGCTCAGCCGCTGATCTCCATCTCTGGACCGG 829  
Db 1458 ACCAAGAAAGGTTTACTCATCTGCTGCTGAGACAAAGCA-----G 1499  
QY 830 ATCAGACTACAGGATCAGTCCTGCTTGAATGCTGGCTGTATCATCTCACCCGCGCTC 889  
Db 1500 GTAAAGCAATTACCGATTTTCCGCTGCGCAACAACTGGTACTCATATTTCCTCCGAGGCTC 1559  
QY 890 ACCTTCCCTCACTCAGGCTGTGTGAGCACTTACTCTGAGCTGGCGGATGACATCTGC 949  
Db 1560 ACCTTCCAGTGCCTGGAGACTGTGTGAACCACTATTCTGAGGTGGCTGTGCTGTGC 1619  
QY 950 TGCCTACTCAAGAGCCCTGTGTCTGAGAGGGCTGGCCGCTCCCTGGCAAGGATATA 1009  
Db 1620 TGTGTGCTCACACGCCCTGCTGACACAAAGCAGCGCTGCCAGCAGTGAGGGCTCC 1679  
QY 1010 CCCCTACCTGTGCTGACAGGAGACCACTCACTGGAAGAGCTGGACAG 1062  
Db 1680 AGCTCACCTGTACCTTGGTCAGAGAGACTGTGACTGGAGAGAGTGTCCAG 1732

## RESULT 9

US-09-864-761-2829  
; Sequence 2829, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761

```

; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2829
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031662.24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
US-09-864-761-2829

Query Match 12.0%; Score 141.8; DB 9; Length 432;
Best Local Similarity 95.4%; Pred. No. 4.7e-33;
Matches 146; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 912 TGGTGGACCAATTACTCTGAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTG 971
Db 253 TGGAGGTCTCTTCTCTAGAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTG 312
QY 972 TCCTGCGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGTCAGA 1031
Db 313 TCCTGCGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGTCAGA 372
QY 1032 GGACACCACTCAACTGGAAGAGCTGGACAGCT 1064
Db 373 GGACACCACTCAACTGGAAGAGCTGGACAGGT 405

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US 09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15513
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031662.24
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
US-09-864-761-15513

Query Match 12.0%; Score 141.8; DB 9; Length 448;
Best Local Similarity 95.4%; Pred. No. 4.7e-33;
Matches 146; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 912 TGGTGGACCAATTACTCTGAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTG 971
Db 269 TGGAGGTCTCTTCTCTAGAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTG 328
QY 972 TCCTGCGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGTCAGA 1031

```

Db 329 TCCTGACAGGGCTGCCCGCTCCCTGGCAAGATATACCTACCTGCTGACTGTGCAGA 388  
 QY 1032 GGACACCACTCACTGGAAGAGCTGGACAGCT 1064  
 Db 389 GGACACCACTCACTGGAAGAGCTGGACAGCT 421

## RESULT 11

US-09-814-353-4631  
 ; Sequence 4631, Application US/09814353  
 ; Publication No. US20030165831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lee, John  
 ; APPLICANT: Thompson, Pamela  
 ; APPLICANT: Lillie, James  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; THERAPY OF OVARIAN CANCER

FILE REFERENCE: MRI-006B  
 CURRENT APPLICATION NUMBER: US/09/814,353  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: US 60/191,031  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: US 60/207,124  
 PRIOR FILING DATE: 2000-05-25  
 PRIOR APPLICATION NUMBER: US 60/211,940  
 PRIOR FILING DATE: 2000-06-15  
 PRIOR APPLICATION NUMBER: US 60/220,661  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: US 60/257,672  
 PRIOR FILING DATE: 2000-12-21  
 NUMBER OF SEQ ID NOS: 22037  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 4631  
 LENGTH: 152  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 NAME/KEY: misc feature  
 LOCATION: 17, 102, 112  
 OTHER INFORMATION: n = A,T,C or G  
 US-09-814-353-4631

Query Match 11.1%; Score 131.6; DB 10; Length 152;  
 Best Local Similarity 95.0%; Pred. No. 4.8e-30;  
 Matches 134; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 536 GGCCCGGCGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGA 595  
 Db 12 GTCGNGCCGAGGTGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGA 71  
 QY 596 GACTGGTGGACGGTGTCTGAGTCTCAGGACAGAGTATACATCCCCACGCTCCAC 655  
 Db 72 GACTGGTGGACGGTGTCTGAGTCTCANGAGAGTNTAACATCCCCACGCTCCAC 131  
 QY 656 GTGGCCAAAGTCTCCCATGGG 676  
 Db 132 GTGGCCAAAGTCTCCCATGGG 152

## RESULT 12

US-09-814-353-10930  
 ; Sequence 10930, Application US/09814353  
 ; Publication No. US20030165831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lee, John  
 ; APPLICANT: Thompson, Pamela  
 ; APPLICANT: Lillie, James  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; THERAPY OF OVARIAN CANCER

FILE REFERENCE: MRI-006B  
 CURRENT APPLICATION NUMBER: US/09/814,353  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: US 60/191,031  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: US 60/207,124  
 PRIOR FILING DATE: 2000-05-25  
 PRIOR APPLICATION NUMBER: US 60/211,940  
 PRIOR FILING DATE: 2000-06-15  
 PRIOR APPLICATION NUMBER: US 60/216,820  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: US 60/220,661  
 PRIOR FILING DATE: 2000-07-25  
 PRIOR APPLICATION NUMBER: US 60/257,672  
 PRIOR FILING DATE: 2000-12-21  
 NUMBER OF SEQ ID NOS: 22037  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 10930  
 LENGTH: 152  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 NAME/KEY: misc feature  
 LOCATION: 17, 102, 112  
 OTHER INFORMATION: n = A,T,C or G  
 US-09-814-353-10930

Query Match 11.1%; Score 131.6; DB 10; Length 152;  
 Best Local Similarity 95.0%; Pred. No. 4.8e-30;  
 Matches 134; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 536 GGCCCGGCGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGA 595  
 Db 12 GTCGNGCCGAGGTGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGA 71  
 QY 596 GACTGGTGGACGGTGTCTGAGTCTCAGGACAGAGTATACATCCCCACGCTCCAC 655  
 Db 72 GACTGGTGGACGGTGTCTGAGTCTCANGAGAGTNTAACATCCCCACGCTCCAC 131  
 QY 656 GTGGCCAAAGTCTCCCATGGG 676  
 Db 132 GTGGCCAAAGTCTCCCATGGG 152

## RESULT 13

US-10-062-674-2038  
 ; Sequence 2038, Application US/10062674  
 ; Publication No. US20040005559A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.  
 ; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
 ; FILE REFERENCE: PA-0026-1 CIP  
 ; CURRENT APPLICATION NUMBER: US/10/062,674  
 ; CURRENT FILING DATE: 2002-01-30  
 ; PRIOR APPLICATION NUMBER: US 09/625,102  
 ; PRIOR FILING DATE: 2000-07-24  
 ; NUMBER OF SEQ ID NOS: 2217  
 ; SOFTWARE: PEEL Program  
 ; SEQ ID NO 2038  
 ; LENGTH: 2343  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20040005559A1 411296.5  
 ; US-10-062-674-2038

Query Match 10.5%; Score 124.6; DB 15; Length 2343;  
 Best Local Similarity 54.1%; Pred. No. 1.3e-27;  
 Matches 284; Conservative 0; Mismatches 229; Indels 12; Gaps 1;  
 QY 457 CCAAGCCAGGACCTGTGACCATGGAGCAGAGAGCAAGCCACGCCGTGGCCCT 516



```

Db 303 CAACAGCAACACACAGGAATCAGGAGGAGGAGGCTCTGAGGACATCATCTGTTGCCCT 362
QY 517 GGGCAGTTTCCGGCAGGTGGCCCGCCAGCTGTGCTGAGACTCGGGAGCCATTGAC 576
Db 363 GTATGATTAAGAGGCCATTTCACACAGAGACCTCAGCTTCCAGAAGGGGGACCCAGATGGT 422
QY 577 CATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGTAAGTCTCAGGCAGAGAGTA 636
Db 423 GGTCTCTAGAGGAATCCGGGGAGTGGTGGAGACTCGATCCCTGGCCACCCGGAGGAGGG 482
QY 637 TAACATCCCCAGGTTCCAGTGGGCAAGTCTCCCAT-----GGGTGGCTGTA 684
Db 483 CTACATCCCAAGCAACTATGTGCGCCGGTGTGACTCTCTGGAGACAGAGGAGTGGTTTTT 542
QY 685 TCAGGGCCTGAGCAGGGAGAAAGCAGAGAACTGTGTGTTTACCTGGGAACCCCTGGAGG 744
Db 543 CAAGGGCATCAGCCGGAAGGACGACAGAGGCCCACTGTGGCTCCCGSCAACATGCTGGG 602
QY 745 GGCCTTCCTCATCCGGGAGAGCCAGACAGAGAGGCTTACTCTGTCTGTCAGTCCGCT 804
Db 603 CTCCTTCATGATCCGGGATAGCGAGACCACCTAAAGGAAGCTACTCTTTGTCCGTGCGAGA 662
QY 805 CAGCCGCCCTGCATCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACATGG 864
Db 663 CTACGACCTCGGAGGAGATACCGTGAACATTACAAGATCCGACCCCTGGACAACGG 722
QY 865 CTGGCTGTACATCTACCGCGCTCACCTTCCCTCACTCCAGGCCCTGGTGGACCATTA 924
Db 723 GGGCTTCTACATATCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCACTA 782
QY 925 CTCTGAGCTGGGGATGACATCTGCTGCTTACTCAAGGAGCCCTG 969
Db 783 CAAGAAGGGAGACGACGGGCTCTGCCAGAAACTGTCTGGTGCCTG 827

```

Search completed: February 20, 2004, 05:59:08  
 Job time : 2861.55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 01:45:20 ; Search time 4572 Seconds  
(without alignments)  
11214.963 Million cell updates/sec

Title: US-09-939-853A-74  
Perfect score: 1183  
Sequence: 1 agctagagctccaaggacc.....tctcttgatgacgcctag 1183

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.man.\*
- 37: em.htg.vrt.\*
- 38: em.ly.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183	100.0	1183	6	AX443133 Sequence
2	1183	100.0	1183	6	AX443135 Sequence
3	1132	95.7	2567	6	AX452880 Sequence
4	1046	88.4	2538	9	BC042041 Homo sapi
5	735	62.1	786	6	AX511153 Sequence
6	735	62.1	786	6	AX512845 Sequence
7	735	62.1	786	9	AF290985 Homo sapi
8	735	62.1	786	9	AF326353 Homo sapi
9	735	62.1	786	9	AK025645 Homo sapi
10	684	57.8	2788	6	AX780857 Sequence
11	483	40.8	737	6	AX511155 Sequence
12	483	40.8	737	9	AP290986 Homo sapi
13	354	29.9	66741	9	HS460J8
14	354	29.9	145833	2	AC026539 Homo sapi
15	152	12.8	145068	9	HSJ97781
16	152	12.8	145833	2	AC026539 Homo sapi
17	43	3.6	777	6	AX511151 Sequence
18	43	3.6	1321	10	AF434390 Mus muscu
19	43	3.6	1348	6	AX511150 Sequence
20	43	3.6	1348	10	AF287467 Mus muscu
21	43	3.6	1384	10	BC052655 Mus muscu
22	41	3.5	163240	10	AL935150 Mouse DNA
23	30	2.5	30	6	AX452884 Sequence
24	29	2.5	299947	2	AC125701 Rattus no
25	29	2.5	322940	2	AC123560 Rattus no
26	26	2.2	26	6	AX443200 Sequence
27	25	2.1	25	6	AX511164 Sequence
28	25	2.1	25	6	AX511165 Sequence
29	25	2.1	25	6	AX511166 Sequence
30	23	1.9	23	6	AX511163 Sequence
31	23	1.9	8028	2	AC014511 Drosophil
32	23	1.9	14560	10	AB011527 Rattus no
33	23	1.9	177577	3	AC022349 Drosophil
34	23	1.9	179016	3	AC018489 Drosophil
35	23	1.9	189516	2	AC128065 Rattus no
36	23	1.9	270729	2	AC133702 Rattus no
37	23	1.9	300469	3	AE003503 Drosophil
38	22	1.9	22	6	AX443201 Sequence
39	22	1.9	475	9	HSPPROT04
40	22	1.9	3070	6	E24027 Human.P pro
41	22	1.9	3070	6	E24028 Method for
42	22	1.9	3070	9	HUMPPRO
43	22	1.9	3082	9	BC012097 Homo sapi
44	22	1.9	67718	2	AC135349 Homo sapi
45	22	1.9	67718	2	AC135349 Homo sapi

ALIGNMENTS

RESULT 1	AX443133	1183 bp	DNA	linear	PAT 02-JUL-2002
AX443133	Sequence 74 from Patent WO0216599.				
LOCUS	AX443133				
DEFINITION	Sequence 74 from Patent WO0216599.				
ACCESSION	AX443133				
VERSION	AX443133.1	GI:21690555			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,				
	Shankets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.B.,				
	Topper,J.N. and Yang,R.B.				

TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES Location/Qualifiers  
source 1..1183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1183; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGCTAGAGCTCAAGGACCCACGCTGTGTCTGTGTGACAGAGCTCAAGGGCCCTGGG 60  
DB 1 AGCTAGAGCTCAAGGACCCACGCTGTGTCTGTGTGACAGAGCTCAAGGGCCCTGGG 60  
QY 61 CTTTCCCTCCCTGGCTCGGCTGTGTGTGGAGGTTTCCCAAGTCCAGATCCCTAAGAG 120  
DB 61 CTTTCCCTCCCTGGCTCGGCTGTGTGTGGAGGTTTCCCAAGTCCAGATCCCTAAGAG 120  
QY 121 CATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180  
DB 121 CATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180  
QY 181 ACCCAACACACCTAGCTTCCCTGAAGATCTCCAGGCTGAGAGGTTCTGGGTG 240  
DB 181 ACCCAACACACCTAGCTTCCCTGAAGATCTCCAGGCTGAGAGGTTCTGGGTG 240  
QY 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTGCCA 300  
DB 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTGCCA 300  
QY 301 GCCAGAGATGGTCTCAGAGAGCTGTCTTCCCAAGCTTTGATGACAAACCAATTTC 360  
DB 301 GCCAGAGATGGTCTCAGAGAGCTGTCTTCCCAAGCTTTGATGACAAACCAATTTC 360  
QY 361 CTCGATGATGTCTCTCAGTGTCTCTGAGGAGCAATGGAGTCTGCCAGCAGAG 420  
DB 361 CTCGATGATGTCTCTCAGTGTCTCTGAGGAGCAATGGAGTCTGCCAGCAGAG 420  
QY 421 AAAATCTGCGAAGCCCAAGCTTGAGTTCCTGTCTCAAGGGCAGGAGCTGTGACCAT 480  
DB 421 AAAATCTGCGAAGCCCAAGCTTGAGTTCCTGTCTCAAGGGCAGGAGCTGTGACCAT 480  
QY 481 GGAAGCAGAGAGAGAGGACAGCCGCTGGCCCTGGCAGTTTCCGCGAGGTGGCC 540  
DB 481 GGAAGCAGAGAGAGAGGACAGCCGCTGGCCCTGGCAGTTTCCGCGAGGTGGCC 540  
QY 541 GCGCGAGCTGTCTGAGATCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGCTG 600  
DB 541 GCGCGAGCTGTCTGAGATCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGCTG 600  
QY 601 GTGACAGGTGTCTGAGTCTCAGCAGAGAGATATAACATCCAGAGGTCCAGCTGGG 660  
DB 601 GTGACAGGTGTCTGAGTCTCAGCAGAGAGATATAACATCCAGAGGTCCAGCTGGG 660  
QY 661 CAAAGTCTCCCATGGTGGTGTATGAGGCTGAGCAGGAGAGAAAGCAGAGAACTGCT 720  
DB 661 CAAAGTCTCCCATGGTGGTGTATGAGGCTGAGCAGGAGAGAAAGCAGAGAACTGCT 720  
QY 721 GTTCTTACCTGGGAACCTCGAGGGGCTTCTCATCCCGGAGAGCCAGACAGGAGAGG 780  
DB 721 GTTCTTACCTGGGAACCTCGAGGGGCTTCTCATCCCGGAGAGCCAGACAGGAGAGG 780  
QY 781 CTCTTACTCTGTGATCCGCTCAGCCGCTTGCATCTCTGGGACCGGATCAGACACTA 840  
DB 781 CTCTTACTCTGTGATCCGCTCAGCCGCTTGCATCTCTGGGACCGGATCAGACACTA 840  
QY 841 CAGATCCATGCTTACATGGTGGTGTGATCTCATCCCGGCTCAGCTTCCCTC 900  
DB 841 CAGATCCATGCTTACATGGTGGTGTGATCTCATCCCGGCTCAGCTTCCCTC 900

QY 901 ACTCCAGCCCTGGTGGACCACTTACTCTGAGCTGGGGATGACATCTGTGCTTACTCAA 960  
DB 901 ACTCCAGCCCTGGTGGACCACTTACTCTGAGCTGGGGATGACATCTGTGCTTACTCAA 960  
QY 961 GGAGCCCTGTCTCTGACAGAGGCTGGCCGCTTCCCTGGCAAGGATATACCCCTAGCTGT 1020  
DB 961 GGAGCCCTGTCTCTGACAGAGGCTGGCCGCTTCCCTGGCAAGGATATACCCCTAGCTGT 1020  
QY 1021 GACTGTGACAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTTCTGA 1080  
DB 1021 GACTGTGACAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTTCTGA 1080  
QY 1081 AGCTGCCACAGGGAGAGTCTTCTCAGTGAAGGTCTCCGGAGTCCCTCAGCTTCTA 1140  
DB 1081 AGCTGCCACAGGGAGAGTCTTCTCAGTGAAGGTCTCCGGAGTCCCTCAGCTTCTA 1140  
QY 1141 CATCAGCCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183  
DB 1141 CATCAGCCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183  
RESULT 2  
AX443135/c  
LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002  
DEFINITION Sequence 76 from Patent WO0216599.  
ACCESSION AX443135  
VERSION AX443135.1 GI:21690556  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,  
Shinkens,R.A., Sytek,K.A., Szekeres,E.S., Tomlinson,J.E.,  
Topper,J.N. and Yang,R.B.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES  
source 1..1183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1183; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGCTAGAGCTCAAGGACCCACGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
DB 1183 AGCTAGAGCTCAAGGACCCACGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGG 1124  
QY 61 CTTTCCCTCCCTGGCTCGGCTGTGTGGAGGTTTCCCAAGTCCAGATCCCTAAGAG 120  
DB 1123 CTTTCCCTCCCTGGCTCGGCTGTGTGGAGGTTTCCCAAGTCCAGATCCCTAAGAG 1064  
QY 121 CATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180  
DB 1063 CATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 1004  
QY 181 ACCCAACCAACACTAGCTTCCCTGAGATCCCTCAGGCTGAGAGGTTCTGGGTG 240  
DB 1003 ACCCAACCAACACTAGCTTCCCTGAGATCCCTCAGGCTGAGAGGTTCTGGGTG 944  
QY 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTGCCA 300  
DB 943 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTGCCA 884  
QY 301 GCCAGAGATGCTGCTCAGCAGAGCTGTCTTCCCAAGCTTTGATGACAAACCAATTTC 360

883 GCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCAAAGCCCTTGTATGACAAACCAATTTCC 824  
 361 CTCATGATGCTGCTCTGAGTGTCTGCTGAGGAAACAATGGGAGTCTCTCCAGCAGAGAG 420  
 823 CTCGATGATGCTGCTCTGAGTGTCTGCTGAGGAAACAATGGGAGTCTCTCCAGCAGAGAG 764  
 421 AAAATCTCTGCCAAGCCCAAGCTGTAGTTCCTCTGTCCTCAAGCCAGGAGCTGTGACCAT 480  
 763 AAAATCTCTGCCAAGCCCAAGCTGTAGTTCCTCTGTCCTCAAGCCAGGAGCTGTGACCAT 704  
 481 GGAAGCAGAGAGAGCAAGCCCAAGCTGTAGTTCCTCTGTCCTCAAGCCAGGAGCTGTGACCAT 480  
 703 GGAAGCAGAGAGAGCAAGCCCAAGCTGTAGTTCCTCTGTCCTCAAGCCAGGAGCTGTGACCAT 704  
 541 GGCAGAGTGTCTGCTGAGACTCGGGAGCCATGACCATGCTCTCTGAGGATGGAGACTG 584  
 643 GGCAGAGTGTCTGCTGAGACTCGGGAGCCATGACCATGCTCTCTGAGGATGGAGACTG 584  
 601 GTGACGGTGTCTGCTGAGTCTCAGGAGAGTATGACCATGCTCTCTGAGGATGGAGACTG 600  
 583 GTGACGGTGTCTGCTGAGTCTCAGGAGAGTATGACCATGCTCTCTGAGGATGGAGACTG 600  
 661 CAAAGTCTCCATGGGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 720  
 523 CAAAGTCTCCATGGGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 720  
 721 GTTGTACTCTGGGAACTCTGAGTCTCAGGAGAGTATGACCATGCTCTCTGAGGATGGAGACTG 780  
 463 GTTGTACTCTGGGAACTCTGAGTCTCAGGAGAGTATGACCATGCTCTCTGAGGATGGAGACTG 780  
 781 CTTTACTCTCTGTCAGTCCGCTCAGCGCCCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 840  
 403 CTTTACTCTCTGTCAGTCCGCTCAGCGCCCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 840  
 841 CAGGATCACTGCTGTGACATGAGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 900  
 343 CAGGATCACTGCTGTGACATGAGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 900  
 901 ACTCAGCGCCCTGTGAGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 960  
 283 ACTCAGCGCCCTGTGAGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 960  
 961 GGAGCCCTGTCTGTCAGAGGGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 1020  
 223 GGAGCCCTGTCTGTCAGAGGGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 1020  
 1021 GACTGTGAGAGGACACCACTCACTGAGAGAGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 1080  
 163 GACTGTGAGAGGACACCACTCACTGAGAGAGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 1080  
 1081 AGCTGCCACAGGAGAGTGTCTTCTCAGTGGGCTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 1140  
 103 AGCTGCCACAGGAGAGTGTCTTCTCAGTGGGCTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 1140  
 1141 CATCAGCTGAATCAGGAGTGTCTTCTCAGTGGGCTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 1183  
 43 CATCAGCTGAATCAGGAGTGTCTTCTCAGTGGGCTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 1183

RESULT 3  
 AX452880  
 LOCUS 2567 bp DNA linear PAT 06-JUL-2002  
 DEFINITION Sequence 1 from Patent WO242457.  
 ACCESSION AX452880  
 VERSION AX452880.1 GI:21712520  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villar, J.J. and  
 Kanner, S.B.

TITLE Cloning and expression of human slap-2: a novel sh2/sh3 domain-containing human slap homologue having immune cell-specific expression  
 JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;  
 Bristol-Myers Squibb Co. (US)  
 FEATUERS Location/Qualifiers  
 source 1..2567  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 95.7%; Score 1132; DB 6; Length 2567;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 AGCTAGAGCTCCAAAGGAGCCAGCCCTGTGTCTGTGACAGAGCTCAAGAGGCGCTGGG 60  
 Db 18 AGCTAGAGCTCCAAAGGAGCCAGCCCTGTGTCTGTGACAGAGCTCAAGAGGCGCTGGG 77  
 Qy 61 CCTTCCCTCCCTGGCTGGCTGTGCTGGAGGGTTCCTCAGTCCAGATCCCTAAGGAG 120  
 Db 78 CCTTCCCTCCCTGGCTGGCTGTGCTGGAGGGTTCCTCAGTCCAGATCCCTAAGGAG 137  
 Qy 121 CATGGGCGAGCTGATCCATCCCTGGTGTACAAACTGCTCACTGCACAGACAGATGCTGAGCT 180  
 Db 138 CATGGGCGAGCTGATCCATCCCTGGTGTACAAACTGCTCACTGCACAGACAGATGCTGAGCT 197  
 Qy 181 ACCCAACCAACACCTAGAGCTCTCCCTGAAGATCTCCAGAGTCCAGAGTTCGGGTG 240  
 Db 198 ACCCAACCAACACCTAGAGCTCTCCCTGAAGATCTCCAGAGTCCAGAGTTCGGGTG 257  
 Qy 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTTAACCTGTCCA 300  
 Db 258 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTTAACCTGTCCA 317  
 Qy 301 GCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCAGAGCTTTGATGACAAACCAATTTCC 360  
 Db 318 GCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCAGAGCTTTGATGACAAACCAATTTCC 377  
 Qy 361 CTCGATGATGCTCTCAGTGTCTCTGAGTGTCTCTGAGGAAACAATGGGAGTCTGCCAGCAGAG 420  
 Db 378 CTCGATGATGCTCTCAGTGTCTCTGAGTGTCTCTGAGGAAACAATGGGAGTCTGCCAGCAGAG 437  
 Qy 421 AAAATCTCTGCCAAGCCCAAGCTGTGAGTTCCTCTGTCCTCAAGCCAGGAGCTGTGACCAT 480  
 Db 438 AAAATCTCTGCCAAGCCCAAGCTGTGAGTTCCTCTGTCCTCAAGCCAGGAGCTGTGACCAT 497  
 Qy 481 GGAAGCAGAGAGAGCAAGCCCAAGCTGTGAGTTCCTCTGTCCTCAAGCCAGGAGCTGTGACCAT 540  
 Db 498 GGAAGCAGAGAGAGCAAGCCCAAGCTGTGAGTTCCTCTGTCCTCAAGCCAGGAGCTGTGACCAT 557  
 Qy 541 GGCAGAGTGTGCTGTGAGACTCGGGAGCCATTTGACCAATCGTCTCTGAGGATGGAGACTG 600  
 Db 558 GGCAGAGTGTGCTGTGAGACTCGGGAGCCATTTGACCAATCGTCTCTGAGGATGGAGACTG 617  
 Qy 601 GTGAGAGTGTGCTGTGAGTCTCAGGAGAGTATTAACATCCCGAGCTCCAGTGGG 660  
 Db 618 GTGAGAGTGTGCTGTGAGTCTCAGGAGAGTATTAACATCCCGAGCTCCAGTGGG 677  
 Qy 661 CAAAGTCTCCATGGGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 720  
 Db 678 CAAAGTCTCCATGGGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAACTGCT 737  
 Qy 721 GTTGTACTCTGGGAACTCTGAGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAGG 780  
 Db 738 GTTGTACTCTGGGAACTCTGAGTGTGCTGTATGAGGGCTGTAGCAGGAGAGAAAGCAGAGGAGG 797  
 Qy 781 CTCCTACTCTCTGTCAGTCCGCTCAGCGCCCTGTGATCTCTGGAGAGCCAGACAGGAGG 840  
 Db 798 CTCCTACTCTCTGTCAGTCCGCTCAGCGCCCTGTGATCTCTGGAGCCGATTCAGACACTA 857  
 Qy 841 CAGGATCCACTGCTTGAACAATGGCTGTGATCATCTCACCGCGCCTCACTTCCCTC 900



Db 858 CAGGATCCATCGCTTGAATGGCTGTGTATCTCACCGCGCTCACCTTCCCCTC 917  
 QY 901 ACTCCAGGCGCTGTGGACCATCTCTGAGCTGGGATGATCTCTGCTGCTACTCAA 960  
 Db 918 ACTCCAGGCGCTGTGGACCATCTCTGAGCTGGGATGATCTCTGCTGCTACTCAA 977  
 QY 961 GGAGCGCTGTGCTGCGAGAGGCTGGCGCGCTCTCCGCAAGGATATACCCCTACCTGT 1020  
 Db 978 GGAGCGCTGTGCTGCGAGAGGCTGGCGCGCTCTCCGCAAGGATATACCCCTACCTGT 1037  
 QY 1021 GACTGTGAGAGGACACCACTCACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGA 1080  
 Db 1038 GACTGTGAGAGGACACCACTCACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGA 1097  
 QY 1081 ACTGCGCAGGAGGAGTCTCTTCTCAGTGAGGCTCTCCGGAGTCCCTCAGCTTCTA 1140  
 Db 1098 AGCTGCCAGGAGGAGTCTCTTCTCAGTGAGGCTCTCCGGAGTCCCTCAGCTTCTA 1157  
 QY 1141 CATCAGCTGAATGACGAGGCTCTCTTCTGATGATCCCTAG 1183  
 Db 1158 CATCAGCTGAATGACGAGGCTCTCTTCTGATGATCCCTAG 1200

RESULT 4  
 BC042041 2538 bp mRNA linear PRI 07-OCT-2003  
 DEFINITION Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (CDNA  
 clone MGC:49845 IMAGE:429896), complete cds.  
 ACCESSION BC042041  
 VERSION 1  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2538)  
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Hopkins, R.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Adkins, R., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Stachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Steeh, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, J.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)

22388257  
 12477932  
 2 (bases 1 to 2538)  
 Strausberg, R.  
 Direct Submission  
 Submitted (23-DEC-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [gcaps-x@mail.nih.gov](mailto:gcaps-x@mail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)

REMARK  
 COMMENT

DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 88 Row: a Column: 20  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 28416422.

FEATURES  
 Location/Qualifiers  
 1..2538  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:49845 IMAGE:4429896"  
 /tissue\_type="Prostate, adenocarcinoma."  
 /clone\_lib="NIH MGC\_91"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORE16"  
 1..2538  
 /gene="SLA2"  
 /note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845,  
 SLAP2"  
 /db\_xref="LocusID:84174"  
 /db\_xref="MIM:606577"  
 363..1148  
 /codon\_start=1  
 /product="Src-like-adaptor 2, isoform a"  
 /protein\_id="AAH42041.1"  
 /db\_xref="GI:27469842"  
 /db\_xref="LocusID:84174"  
 /translation="MGSLLPSRRKSLPSLSVSSVQGGPVTMEATSKATAVALGSPFP  
 AGPAELSLRIGELPTIVSEGDWVLSVSEGREYNIPIVHVAKVSHGWIYEGLSRE  
 KARELLPLGNPGAFILRESQTRGYSLSVLSRSPASWDRIHRIHLDNGLWLYI  
 SPRTTFPLQALVDHYSELADICLLKPCVLRQAGLPKQDPLFVTVQRTPLNWK  
 ELDSLLFSEAAATGEESSLSEGLRESLSFYISLNDKAVSLDDA"  
 468..632  
 /note="SH3; Region: SH3 domain. SH3 (Src homology 3)  
 domains are often indicative of a protein involved in  
 signal transduction related to cytoskeletal organization.  
 First described in the Src cytoplasmic tyrosine kinase.  
 The structure is a partly opened beta barrel."  
 /db\_xref="CDD:pfam00018"  
 642..890  
 /note="SH2; Region: SH2 domain"  
 /db\_xref="CDD:pfam00017"

gene  
 CDS  
 misc\_feature  
 misc\_feature

ORIGIN  
 Query Match 88.4%; Score 1046; DB 9; Length 2538;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 36 GTGACAGAGCTCAAAGGGCCCTGGGCGCTTCCTCCCTGGCTGGCTGTGCTTGGAGGGT 95  
 Db 1 GTGACAGAGCTCAAAGGGCCCTGGGCGCTTCCTCCCTGGCTGGCTGTGCTTGGAGGGT 60  
 QY 96 TCCCGAGTCCAGATCCCTTAGGAGCATGGGCGAGCTGATCCATCCCTGGTGTACAAACT 155  
 Db 61 TCCCGAGTCCAGATCCCTTAGGAGCATGGGCGAGCTGATCCATCCCTGGTGTACAAACT 120  
 QY 156 GCTGACTGCAGACAGATGCTGAGCTACCCAAACCAACACCTAGCTCTCTCCCTGAAGATCC 215  
 Db 121 GCTGACTGCAGACAGATGCTGAGCTACCCAAACCAACACCTAGCTCTCTCCCTGAAGATCC 180  
 QY 216 TCCCGAGTCCAGATGCTGAGCTACCCAAACCAACACCTAGCTCTCTCCCTGAAGATCC 275  
 Db 181 TCCCGAGTCCAGATGCTGAGCTACCCAAACCAACACCTAGCTCTCTCCCTGAAGATCC 240  
 QY 276 GGCCCCCAAGCCCTTAACCTGTCCAGCCAGAGATGGCTCTCAGCAGAGCTGCTTCCCA 335

Db 241 GGCCCCAAGCCCTAACCTGCTCCAGCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCA 300  
 QY 336 AGCTTTTGATGACAAACCAATTTCCCTGATGATGCTTCTGAGTGTCTGCTGAGGAA 395  
 Db 301 AGCTTTTGATGACAAACCAATTTCCCTGATGATGCTTCTGAGTGTCTGCTGAGGAA 360  
 QY 396 CAATGGGAAGTCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCCTCTG 455  
 Db 361 CAATGGGAAGTCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCCTCTG 420  
 QY 456 TCCAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACACGCCCTGGGCC 515  
 Db 421 TCCAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACACGCCCTGGGCC 480  
 QY 516 TGGGCAATTTCCCGCAGTGTGCGCGCGAGCTGTGCTGAGAGCTGGGAGCCATGA 575  
 Db 481 TGGGCAATTTCCCGCAGTGTGCGCGCGAGCTGTGCTGAGAGCTGGGAGCCATGA 540  
 QY 576 CCATCTCTCTGAGGATGAGACTGTGTGACGCTGTCTCTGAACTCTCAGGCAGAGAGT 635  
 Db 541 CCATCTCTCTGAGGATGAGACTGTGTGACGCTGTCTCTGAACTCTCAGGCAGAGAGT 600  
 QY 636 ATAACTATCCCGCAGTGTGCGCGCGAGCTGTGCTGAGAGCTGGGAGCCATGA 695  
 Db 601 ATAACTATCCCGCAGTGTGCGCGCGAGCTGTGCTGAGAGCTGGGAGCCATGA 660  
 QY 696 GCAGGAGAGAAAGCAGAGGAATCTGTGTTTACCTGGGAAACCTGGAGGGGCTTCTCA 755  
 Db 661 GCAGGAGAGAAAGCAGAGGAATCTGTGTTTACCTGGGAAACCTGGAGGGGCTTCTCA 720  
 QY 756 TCGGAGAGCCAGACAGGAGGCTCTTACTCTCTGCTGAGTCCGCTCAGCGCCCTG 815  
 Db 721 TCGGAGAGCCAGACAGGAGGCTCTTACTCTCTGCTGAGTCCGCTCAGCGCCCTG 780  
 QY 816 CATCTCTGGACCGGATCAGACACTACAGGATCCACTGCTTGAATGGCTGGCTGTACA 875  
 Db 781 CATCTCTGGACCGGATCAGACACTACAGGATCCACTGCTTGAATGGCTGGCTGTACA 840  
 QY 876 TCTCAGCGGCTCAGCTTCCCTCAGTCCAGGCTGCTGAGGAGGCTGAGGAGGCTGAGTGG 935  
 Db 841 TCTCAGCGGCTCAGCTTCCCTCAGTCCAGGCTGCTGAGGAGGCTGAGGAGGCTGAGTGG 900  
 QY 936 CGGATGACATCTGCTGCTTCAAGGAGCCCTGTGCTGCTGAGAGGCTGCGCCGCTCC 995  
 Db 901 CGGATGACATCTGCTGCTTCAAGGAGCCCTGTGCTGCTGAGAGGCTGCGCCGCTCC 960  
 QY 996 CTGGCAGGATATACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055  
 Db 961 CTGGCAGGATATACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
 QY 1056 TGGACAGCTCCCTCTGTTTCTGAGGCTGCCACAGGGGAGGAGTCTTCTCAGTGAGG 1115  
 Db 1021 TGGACAGCTCCCTCTGTTTCTGAGGCTGCCACAGGGGAGGAGTCTTCTCAGTGAGG 1080  
 QY 1116 GTCTCGGAGGCTTCTCAGCTTCTCAGCTTCTCAGCTTCTCAGCTTCTCAGCTTCTCAG 1175  
 Db 1081 GTCTCGGAGGCTTCTCAGCTTCTCAGCTTCTCAGCTTCTCAGCTTCTCAGCTTCTCAG 1140  
 QY 1176 ATGCTTAG 1183  
 Db 1141 ATGCTTAG 1148

RESULT 5  
 AX511153  
 LOCUS AX511153  
 DEFINITION Sequence 4 from Patent WO0242452.  
 ACCESSION AX511153  
 VERSION AX511153.1 GI:23392046  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1  
 AUTHORS McElade, J.C. and Loreto, M.P.  
 TITLE Adapter gene  
 JOURNAL Patent: WO 0242452-A 4 30-MAY-2002;  
 The Hospital for Sick Children (CA)  
 FEATURES  
 source  
 1..786  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 62.1%; Score 735; DB 6; Length 786;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 398 ATGGGAAGTCTGCTCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCCTCTGTC 457  
 Db 1 ATGGGAAGTCTGCTCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCCTCTGTC 60  
 QY 458 CAGGCGAGGACCTGTGACCATGGAAGCAGAGAAAGCCACAGCGGTGGCCCTG 517  
 Db 61 CAGGCGAGGACCTGTGACCATGGAAGCAGAGAAAGCCACAGCGGTGGCCCTG 120  
 QY 518 GGCAGTTCCTCCGCGAGGTGGCCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 577  
 Db 121 GGCAGTTCCTCCGCGAGGTGGCCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180  
 QY 578 ATCGTCTCTGAGGATGGAGACTGTGAGCGGTGTCTGAGTCTCAGGCAGAGAGTAT 637  
 Db 181 ATCGTCTCTGAGGATGGAGACTGTGAGCGGTGTCTGAGTCTCAGGCAGAGAGTAT 240  
 QY 638 AACATCCCCAGCGTCCAGTGGGCAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 697  
 Db 241 AACATCCCCAGCGTCCAGTGGGCAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 300  
 QY 698 AGGGAAGCAGAGGAGTCTGTTTACTTGGGAAACCTGGAGGGGCTTCTCTCATC 757  
 Db 301 AGGGAAGCAGAGGAGTCTGTTTACTTGGGAAACCTGGAGGGGCTTCTCTCATC 360  
 QY 758 CGGAGAGCCAGCAGCAGGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGCA 817  
 Db 361 CGGAGAGCCAGCAGCAGGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGCA 420  
 QY 818 TCCTGGGACCGGATCAGACACTCAGGATCCACTGCTTGCATGCTGCTGCTGCTGCTGCTG 877  
 Db 421 TCCTGGGACCGGATCAGACACTCAGGATCCACTGCTTGCATGCTGCTGCTGCTGCTGCTG 480  
 QY 878 TCACCGCGCTCAGCTTCCCTCTCACTCCAGGCGCTGTGGACCAATTAATCTGAGCTGGCG 937  
 Db 481 TCACCGCGCTCAGCTTCCCTCTCACTCCAGGCGCTGTGGACCAATTAATCTGAGCTGGCG 540  
 QY 938 GATGACATCTGCTGCTTCAAGAGCGCTGTCTGCTGAGAGGCTGCGCGCTCCCT 997  
 Db 541 GATGACATCTGCTGCTTCAAGAGCGCTGTCTGCTGAGAGGCTGCGCGCTCCCT 600  
 QY 998 GCAAGGATATACCCCTACCTGTGACTGTGAGAGGACCACTCAACTGGAAGAGCTG 1057  
 Db 601 GCAAGGATATACCCCTACCTGTGACTGTGAGAGGACCACTCAACTGGAAGAGCTG 660  
 QY 1058 GACAGTCCCTCTGTTTCTGAGAGCTGCGACAGGGGAGGAGTCTCTCTCAGTGAGGGT 1117  
 Db 661 GACAGTCCCTCTGTTTCTGAGAGCTGCGACAGGGGAGGAGTCTCTCTCAGTGAGGGT 720  
 QY 1118 CTCGGGAGTCCCTCAGCTTCTATCAGCGCTGATGACGAGGCTGTCTCTTTGGATGAT 1177  
 Db 721 CTCGGGAGTCCCTCAGCTTCTATCAGCGCTGATGACGAGGCTGTCTCTTTGGATGAT 780  
 QY 1178 GCCTAG 1183  
 Db 781 GCCTAG 786

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

RESULT 6
AX572845
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS
ORIGIN
Query Match
Best Local Similarity
Matches
398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAGTCTCTCTGTC 457
DB
1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAGTCTCTCTGTC 60
458 CAAGCCAGCAGCACTGTGACCATGCAAGCAGAGAGCAAGCAAGCCAGCCGTGGCCCTG 517
DB
61 CAAGCCAGCAGCACTGTGACCATGCAAGCAGAGAGCAAGCAAGCCAGCCGTGGCCCTG 120
459 GGCAGTTTCCCGCAGAGTGGCCCGCCGCGAGCTGTCGTGAGACTCGGGGAGCCATTGACC 577
DB
121 GGCAGTTTCCCGCAGAGTGGCCCGCCGCGAGCTGTCGTGAGACTCGGGGAGCCATTGACC 180
459 ATCTGCTCTGAGGATGGAGACTGCTGTCGACCGGTGCTGTCGAGTCTCAGCAGAGAGTAT 637
DB
181 ATCTGCTCTGAGGATGGAGACTGCTGTCGACCGGTGCTGTCGAGTCTCAGCAGAGAGTAT 240
459 AACATCCCGCAGCTCCAGCTGGGCAAGTCTCCATGGGTGGTGTATGAGGCGCTGAGC 697
DB
241 AACATCCCGCAGCTCCAGCTGGGCAAGTCTCCATGGGTGGTGTATGAGGCGCTGAGC 300
459 AGGAGAGAAAGCAGAGGAAGTCTCTGTTGTTACCTGGGAACCTCGAGGGGCTTCTTCATC 757
DB
301 AGGAGAGAAAGCAGAGGAAGTCTCTGTTGTTACCTGGGAACCTCGAGGGGCTTCTTCATC 360
459 CGGAGAGCCAGCAGAGAGGCTCTTACTCTGTGTCAGTCCGCTCAGCCGCCCTGCA 817
DB
361 CGGAGAGCCAGCAGAGAGGCTCTTACTCTGTGTCAGTCCGCTCAGCCGCCCTGCA 420
459 TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGCATGAGTGGTGGTGTACATC 877
DB
421 TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGCATGAGTGGTGGTGTACATC 480

```

```

878 TCACCGGCGCTCACCTTCCCTCACTCCAGGCGCTGGTGGACCATTTACTCTGAGTGGCG 937
481 TCACCGGCGCTCACCTTCCCTCACTCCAGGCGCTGGTGGACCATTTACTCTGAGTGGCG 540
938 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGCTTCGAGAGGCTGGCCGCTCCCT 997
541 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGCTTCGAGAGGCTGGCCGCTCCCT 600
998 GGCAAGGATATACCCCTACCTGCTGAGCTGTGAGAGGACACCACTCAACTGGAAAGAGCTG 1057
601 GGCAAGGATATACCCCTACCTGCTGAGCTGTGAGAGGACACCACTCAACTGGAAAGAGCTG 660
1058 GAGAGTCTCCCTCTGTTTCTGAGCTGCCAGAGGAGGAGTCTCTCTCAGTCAAGGT 1117
661 GAGAGTCTCCCTCTGTTTCTGAGCTGCCAGAGGAGGAGTCTCTCTCAGTCAAGGT 720
1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTTGAATGAGAGGCTGTCTCTTGGATGAT 1177
721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTTGAATGAGAGGCTGTCTCTTGGATGAT 780
1178 GCCTAG 1183
781 GCCTAG 786

RESULT 7
AF290985
LOCUS
DEFINITION Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS
ORIGIN
Query Match
Best Local Similarity
Matches
62.1%; Score 735; DB 9; Length 786;
99.9%; Pred. No. 0;
785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

TITLE	Payan,D.G., Mancebo,H.S.Y. and Wu,J.
JOURNAL	Functional Cloning of Src-like Adapter
MEDLINE	Inhibitor of Antigen Receptor Signaling
PUBMED	J. Exp. Med. 194 (9), 1263-1276 (2001)
REFERENCE	21553259
AUTHORS	11696592
TITLE	2 (bases 1 to 786)
JOURNAL	Holland,S.J., Mendenhall,M.K., Zhou,X., Spencer,C., Pardo,J.,
FEATURES	Fu,A.C., Sheng,N., Shen,M., Liao,C., Luo,Y., Payan,D.G.,
source	Mancebo,H.S.Y. and Wu,J.
	Direct Submission
	Submitted (05-DEC-2000) Rigel Pharmaceutical Inc., 240 East Grand
	Avenue, South San Francisco, CA 94080, USA
	Location/Qualifiers
	1..786
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosomes="20"
CDS	1..786
	/notes="SLAP2; SH2 and SH3 domain-containing adaptor"
	/codon_start=1
	/product="Src-like adapter protein-2"
	/protein_id="AAI29204.1"
	/db_xref="GI:16797892"
	/translation="MGSLPSRRKSLPSLSVQGGPVTMEARSKATAVALGSPF
	AGQAPLSRLGLEPTIVSEDDQNMVTLSEVSGREYNISVHVAKYSHGMLYEGLSRE
	KAEALLLPNGPQAFLEIRSGTSGYSLVLSRPSAMDRIHRHYRIHCLDNGWLYI
	SPRLTFPSLQALVDHYSLADIDICLLKEPCVLQRAPIPGKDIPLPVTVQRTPLNWK
	ELDSSLPSREATGTGESLSLSEGLRESLSFVLSLNDENAVSLDDA"
ORIGIN	
	Query Match 62.1%; Score 735; DB 9; Length 786;
	Best Local Similarity 99.9%; Pred.No.0;
	Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	398 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAACCCCAAGCTTGAAGTCTCTCTGTC 457
Db	1 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAACCCCAAGCTTGAAGTCTCTCTGTC 60
QY	458 CAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAGGCGACAGCGTGGCCCTG 517
Db	61 CAGGCCAGGACCTGTGTACCATGGAAGCAGAGAGAAGCAGGCGACAGCGTGGCCCTG 120
QY	518 GGCAGATTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTTCGGGGAGCCATTGACC 577
Db	121 GGCAGATTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTTCGGGGAGCCATTGACC 180
QY	578 ATCGTCTCTGAGATGGAGACTGTGGACGCTGTCTGAGCTTCAGCAGAGAGTAT 637
Db	181 ATCGTCTCTGAGATGGAGACTGTGGACGCTGTCTGAGCTTCAGCAGAGAGTAT 240
QY	638 AACATCCCAGCGTCCACGTGGGCAAACTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 697
Db	241 AACATCCCAGCGTCCACGTGGGCAAACTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300
QY	698 AGGAGAAAGCAGAGAACTGCTGTGTGTACTCTGGGAACCTCGAGGGCCCTTCCTCATC 757
Db	301 AGGAGAAAGCAGAGAACTGCTGTGTGTGTACTCTGGGAACCTCGAGGGCCCTTCCTCATC 360
QY	758 CGGAGAGCCAGACAGGAGGCTCTTACTCTCTGTCACTCCGCGCTTCAGCGCCCTGCA 817
Db	361 CGGAGAGCCAGACAGGAGGCTCTTACTCTCTGTCACTCCGCGCTTCAGCGCCCTGCA 420
QY	818 TCTGTGGACCGGATCAGACACTACAGGATCCACTGCCTTCACAAATGGCTGGCTGTACATC 877
Db	421 TCTGTGGACCGGATCAGACACTACAGGATCCACTGCCTTCACAAATGGCTGGCTGTACATC 480
QY	878 TCACCGCGCTCACCTTCCCTCTACTCCAGCCCTGTGTGACCATTAATCTGTAGCTGGCG 937
Db	481 TCACCGCGCTCACCTTCCCTCTACTCCAGCCCTGTGTGACCATTAATCTGTAGCTGGCG 540
QY	938 GATGACATCTGCTCTACTCAAGAGAGCCCTGTCTCTCTGACAGAGGCTGGCCCGCTCCCT 997

TITLE	Payan,D.G., Mancebo,H.S.Y. and Wu,J.
JOURNAL	Functional Cloning of Src-like Adapter
MEDLINE	Inhibitor of Antigen Receptor Signaling
PUBMED	J. Exp. Med. 194 (9), 1263-1276 (2001)
REFERENCE	21553259
AUTHORS	11696592
TITLE	2 (bases 1 to 786)
JOURNAL	Holland,S.J., Mendenhall,M.K., Zhou,X., Spencer,C., Pardo,J.,
FEATURES	Fu,A.C., Sheng,N., Shen,M., Liao,C., Luo,Y., Payan,D.G.,
source	Mancebo,H.S.Y. and Wu,J.
	Direct Submission
	Submitted (05-DEC-2000) Rigel Pharmaceutical Inc., 240 East Grand
	Avenue, South San Francisco, CA 94080, USA
	Location/Qualifiers
	1..786
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosomes="20"
CDS	1..786
	/notes="SLAP2; SH2 and SH3 domain-containing adaptor"
	/codon_start=1
	/product="Src-like adapter protein-2"
	/protein_id="AAI29204.1"
	/db_xref="GI:16797892"
	/translation="WGSIPSRKSLPSLSSVQGGPVTMEARSKATAVALGSPF
	AGQAPLSRLGLEPTIVSEDDQNMVTLSEVSGREYNISVHVAKYSHGMLYEGLSRE
	KAEALLLPNGPQAFLEIRSGTSGYSLVLSRPSAMDRIHRHYRIHCLDNGWLYI
	SPRLTFPSLQALVDHYSLADIDICLLKEPCVLQRAPIPGKDIPLPVTVQRTPLNWK
	ELDSSLPSREATGTGESLSLSEGLRESLSFVLSLNDENAVSLDDA"
ORIGIN	
	Query Match 62.1%; Score 735; DB 9; Length 786;
	Best Local Similarity 99.9%; Pred.No.0;
	Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	398 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAACCCCAAGCTTGAGTTCCTCTGTC 457
Db	1 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAACCCCAAGCTTGAGTTCCTCTGTC 60
QY	458 CAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCACAGCCGTGGCCCTG 517
Db	61 CAGGCCAGGACCTGTGTACCATGGAAGCAGAGAGAAGCAAGGCACAGCCGTGGCCCTG 120
QY	518 GGCAGATTTCCTGGCAGGTGGCCCGCCGAGCTGTGCTGAGACTTCGGGGAGCCATTGACC 577
Db	121 GGCAGATTTCCTGGCAGGTGGCCCGCCGAGCTGTGCTGAGACTTCGGGGAGCCATTGACC 180
QY	578 ATCGTCTCTGAGATGGAGACTGTGGACGTGTCTGTCTGAAGTCTCAGCAGAGAGTAT 637
Db	181 ATCGTCTCTGAGATGGAGACTGTGGACGTGTCTGTCTGAAGTCTCAGCAGAGAGTAT 240
QY	638 AACATCCCAGCGTCCACGTGGGCAAACTCTCCCATGGGTGGCTGTATGAGGCGCTTGAGC 697
Db	241 AACATCCCAGCGTCCACGTGGGCAAACTCTCCCATGGGTGGCTGTATGAGGCGCTTGAGC 300
QY	698 AGGAGAAAGCAGAGAACTGCTGTGTGTACTCTGGGAACCTCGAGGGCCCTTCCTCATC 757
Db	301 AGGAGAAAGCAGAGAACTGCTGTGTGTGTACTCTGGGAACCTCGAGGGCCCTTCCTCATC 360
QY	758 CGGAGAGCCAGACAGGAGGCTCTTACTCTCTGTCACTCCGCGCTTCAGCCGCCCTGCA 817
Db	361 CGGAGAGCCAGACAGGAGGCTCTTACTCTCTGTCACTCCGCGCTTCAGCCGCCCTGCA 420
QY	818 TCTGTGGACCGGATCAGACACTACAGGATCCACTGCCTTCACAAATGGCTGGCTGTACATC 877
Db	421 TCTGTGGACCGGATCAGACACTACAGATCCACTGCCTTCACAAATGGCTGGCTGTACATC 480
QY	878 TCACCGCGCTCACCTTCCCTCTACTCCAGCCCTGTGTGACCATTACTCTGAGCTGGCG 937
Db	481 TCACCGCGCTCACCTTCCCTCTACTCCAGCCCTGTGTGACCATTACTCTGAGCTGGCG 540
QY	938 GATGACATCTGCTCTACTCAAGAGAGCCCTGTCTCTCTGACAGAGGCTGGCCCGCTCCCT 997

```
Db 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTCTCTGTCAGAGGGCTGCGCCGCTCCCT 600
QY 998 GCGAGGATATACCCCTACCTGTGCTGCTGTCAGAGGACACCACTCACTGGAAGAGCTG 1057
Db 601 GCGAGGATATACCCCTACCTGTGCTGCTGTCAGAGGACACCACTCACTGGAAGAGCTG 660
QY 1058 GACAGCTCCCTCCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGGT 1117
Db 661 GACAGCTCCCTCCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGGT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATCAGAGCTGCTCTTCTTGGATGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATCAGAGCTGCTCTTCTTGGATGAT 780
QY 1178 GCCTAG 1183
Db 781 GCCTAG 786

RESULT 9
AK025645
LOCUS Homo sapiens cDNA: FL221992 fis, clone HEP06554. linear PRI 13-SEP-2003
DEFINITION AK025645
ACCESSION AK025645
VERSION AK025645.1 GI:10438227
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2415)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:fidna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'- and one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source Location/Qualifiers
1..2415
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP06554"
/cell_line="HepG2"
/clone_lib="HEP"
/clone="cloning vector pME185FL3"
61..846
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15201.1"
/db_xref="GI:10438228"
/tranlation="MGSLSRRKSLPSLSVSGQGPVMEERSKATAVALGSP
AGGPAELSLRLGEPLTIYSEGDWTVLSEVSGREYNIPSVHVAKVSHGLYGLSRE
KAEELLLPGNPGGAFLLRESQTRGSVLSVRLSRFASMDRIHRIHCLDNGWLYI
SPRLTFPSQLALVHYSLEADICLLKEPCVLRAGPLPGKDIPLPVTQRTPLNWK
ELDSLLFPSEATGEESLLSGLRESUSFYISLNDKAVSLDDA"
```

```
ORIGIN
Query Match 62.1%; Score 735; DB 9; Length 2415;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 398 ATGGCAAGTCTGCCCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTCAGTTCTCTCTGTC 457
Db 61 ATGGCAAGTCTGCCCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTCAGTTCTCTCTGTC 120
QY 458 CAAAGCCAGGAGCCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGCCCCTG 517
Db 121 CAAAGCCAGGAGCCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGCCCCTG 180
QY 518 GGCAGTTTCCCGGAGGTTGGCCCGCCGAGCTGCTGCTGAGACTCCGGGAGCCATTGACC 577
Db 181 GGCAGTTTCCCGGAGGTTGGCCCGCCGAGCTGCTGCTGAGACTCCGGGAGCCATTGACC 240
QY 578 ATCGTCTCTGAGGATGAGACTGCTGTCAGCGTGTCTGTAAGTCTCAGGCAGAGAGTAT 637
Db 241 ATCGTCTCTGAGGATGAGACTGCTGTCAGCGTGTCTGTAAGTCTCAGGCAGAGAGTAT 300
QY 638 AACATCCCGCAGCTCCACGTGGGCAAAAGTCTCCCATCGGTGGCTGTATGAGGGCCTGAGC 697
Db 301 AACATCCCGCAGCTCCACGTGGGCAAAAGTCTCCCATCGGTGGCTGTATGAGGGCCTGAGC 360
QY 698 AGGAGAAAGCAGAGGAACCTGCTGTTTACCTGGGAAACCTTGAGGGGCTTCTCTCATC 757
Db 361 AGGAGAAAGCAGAGGAACCTGCTGTTTACCTGGGAAACCTTGAGGGGCTTCTCTCATC 420
QY 758 CGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCTCGCA 817
Db 421 CGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCTCGCA 480
QY 818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCACATGGCTGGCTGTACATC 877
Db 481 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCACATGGCTGGCTGTACATC 540
QY 878 TCACCGGCTCACCCTTCCCTCCTCAGGAGCCCTGCTGCTGTCAGAGGCTGTCGAGCTGGCG 937
Db 541 TCACCGGCTCACCCTTCCCTCCTCAGGAGCCCTGCTGCTGTCAGAGGCTGTCGAGCTGGCG 600
QY 938 GATGACATCTGCTGCTTACTCAAGAGGAGCCCTGCTGCTGTCAGAGGCTGTCGAGCTGGCG 997
Db 601 GATGACATCTGCTGCTTACTCAAGAGGAGCCCTGCTGCTGTCAGAGGCTGTCGAGCTGGCG 660
QY 998 GGCAGAGTATACCCCTACCTGTCAGAGGAGACCACTCACTCACTGGAAGAGCTG 1057
Db 661 GGCAGAGTATACCCCTACCTGTCAGAGGAGACCACTCACTCACTGGAAGAGCTG 720
QY 1058 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTCTCAGTGAGGGT 1117
Db 721 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTCTCAGTGAGGGT 780
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATCAGAGGCTGCTCTTCTTGGATGAT 1177
Db 781 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATCAGAGGCTGCTCTTCTTGGATGAT 840
QY 1178 GCCTAG 1183
Db 841 GCCTAG 846

RESULT 10
AX780857
LOCUS Homo sapiens 2788 bp DNA linear PAT 14-JUL-2003
DEFINITION Sequence 3014 from Patent WO03039443.
ACCESSION AX780857
VERSION AX780857.1 GI:32697851
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

LOCUS	AF2909386	737 bp	linear	PRI 21-JAN-2000
DEFINITION	Homo sapiens Src-like adaptor protein-2 splice isoform mRNA, complete cds; alternatively spliced.			



[illegible]





```

* 76825 81179: contig of 4355 bp in length
* 81180 81279: gap of 100 bp
* 81280 87009: contig of 5730 bp in length
* 87010 87109: gap of 100 bp
* 87110 90855: contig of 3746 bp in length
* 90856 90955: gap of 100 bp
* 90956 98220: contig of 5865 bp in length
* 98221 98220: gap of 100 bp
* 98221 102321: contig of 5701 bp in length
* 102322 102421: gap of 100 bp
* 102422 108293: contig of 5872 bp in length
* 108294 108393: gap of 100 bp
* 108394 116689: contig of 8296 bp in length
* 116690 116789: gap of 100 bp
* 116790 125364: contig of 8475 bp in length
* 125365 125364: gap of 100 bp
* 125365 136354: contig of 10990 bp in length
* 136355 136454: gap of 100 bp
* 136455 145833: contig of 9379 bp in length.
FEATURES
    source
        1..145833
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /map="20"
            /clone="RP11-712N14"
            /clone_lib="RPC1-11 Human Male BAC"
            1..1214
                /note="assembly_fragment"
            1315..2673
                /note="assembly_fragment"
            2774..4520
                /note="assembly_fragment"
            4621..5961
                /note="assembly_fragment"
            6062..7719
                /note="assembly_fragment"
            7820..9799
                /note="assembly_fragment"
            9900..11434
                /note="assembly_fragment"
            11535..14382
                /note="assembly_fragment"
            14483..17148
                /note="assembly_fragment"
            17249..19239
                /note="assembly_fragment"
            19340..21102
                /note="assembly_fragment"
            21203..23371
                /note="assembly_fragment"
            23472..25782
                /note="assembly_fragment"
            25883..28824
                /note="assembly_fragment"
            28925..31619
                /note="assembly_fragment"
            31720..34679
                /note="assembly_fragment"
            34780..37527
                /note="assembly_fragment"
            37628..40146
                /note="assembly_fragment"
            40247..43743
                /note="assembly_fragment"
            43844..47133
                /note="assembly_fragment"
            47234..51023
                /note="assembly_fragment"
            51124..54935
                /note="assembly_fragment
clone_end:17

```

```

misc_feature      vector_side:right
55036..59553
Query Match
Best Local Similarity 29.9%; Score 354; DB 2; Length 145833;
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCTAGAGCTCCAAGGACCCACGCTGTGTCTGTGTGACAGAGCTCAAGGCCCTGGG 60
DB 99745 AGCTAGAGCTCCAAGGACCCACGCTGTGTCTGTGTGACAGAGCTCAAGGCCCTGGG 99804
QY 61 CTTTCCTCCCTGGCTGGCTGTGTGTGGAGGTTCCCGAGTCCAGTCCAGTCCCTAAGGAG 120
DB 99805 CTTTCCTCCCTGGCTGGCTGTGTGTGGAGGTTCCCGAGTCCAGTCCCTAAGGAG 99864
QY 121 CATGGGGCAGCTGATCCATCCCTGGTGTACAACATGCTGTGCTGACACAGATCTGAGCT 180
DB 99865 CATGGGGCAGCTGATCCATCCCTGGTGTACAACATGCTGTGCTGACACAGATCTGAGCT 99924
QY 181 ACCCAAAACCAACACTAGCTCTCCCTGGAAGATCTCTCCAGGCTGAGAGAGTTCTGGTG 240
DB 99925 ACCCAAAACCAACACTAGCTCTCCCTGGAAGATCTCTCCAGGCTGAGAGAGTTCTGGTG 99984
QY 241 TCCTAGGACCAAGGACACTGGCAGCTTCAGAGGCGCCCAAGGCCCTTAACCTGTCCA 300
DB 99985 TCCTAGGACCAAGGACACTGGCAGCTTCAGAGGCGCCCAAGGCCCTTAACCTGTCCA 100044
QY 301 GCCAGAGCATCGTCTCAGCAGAGCTCTCTTCCCAAGCCTTTGATGACAAACCA 354
DB 100045 GCCAGAGCATCGTCTCAGCAGAGCTCTCTTCCCAAGCCTTTGATGACAAACCA 100098
RESULT 15
HSDJ977B1/c
LOCUS
DEFINITION
    Human DNA sequence from clone RPs-977B1 on chromosome 20. Contains
    ESTs, STSs, GSSs and three putative CpG islands. Contains the 3'
    end of the gene for a novel protein tyrosine kinase, a gene for
    heterogeneous nuclear ribonucleoprotein A3 pseudogene, the gene for
    RAB5-interacting protein, the TGIF2 gene for TGF(beta)-induced
    transcription factor 2 with two isoforms, the MYRL2 gene for myosin
    regulatory light chain 2 (smooth muscle isoform), the 3' end of the
    gene KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4)
    with two isoforms and a novel gene, complete sequence.
ACCESSION
    AL050318
VERSION
    AL050318.13 GI:9581785
KEYWORDS
    HTG; CpG island; heterogeneous ribonucleoprotein; KIAA0964; myosin
    domain; TGIF2; transcription factor; tyrosine kinase.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 145068)
    Lloydi.D.
    Direct Submission
    Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
    CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
    requests: clonerequest@sanger.ac.uk
    On Jul 28, 2000 this sequence version replaced gi:5924017.
    During sequence assembly data is compared from overlapping clones.
    Where differences are found these are annotated as variations.
    Together with a note of the overlapping clone name. Note that the
    variation annotation may not be found in the sequence submission
    corresponding to the overlapping clone, as we submit sequences with
    only a small overlap as described above.
    The following abbreviations are used to associate primary accession
    numbers given in the feature table with their source databases:
    Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
    on the WORMPEP database can be found at
    http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
    was generated from part of bacterial clone contigs of human

```

chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence is the entire insert of clone RP5-977B1. The true left end of clone RP2-460J8 is at 62245 in this sequence. The true right end of clone CDP-218229 is at 62218 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-977B1 is from the library RPI-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.  
Location/Qualifiers

```

1. .145068
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP5-977B1"
/clone_lib="RPI-5"
568. .875
/feature="AlusX repeat: matches 1. .311 of consensus"
1024. .1580
/feature="match: GSS: Em:AQ620498"
1047. .1629
/feature="match: GSS: Em:AQ779162"
1631. .1923
/feature="AlusX repeat: matches 1. .294 of consensus"
2776. .2842
/feature="L2 repeat: matches 2357. .2423 of consensus"
3891. .3949
/feature="MIR repeat: matches 34. .90 of consensus"
4458. .4611
/feature="MIS repeat: matches 67. .231 of consensus"
5133. .5549
/feature="L1MC/D repeat: matches 5432. .5528 of consensus"
5580. .5680
/feature="L1MC/D repeat: matches 5924. .6340 of consensus"
5703. .5890
/feature="FLAM_C repeat: matches 1. .132 of consensus"
5923. .6012
/feature="AlusX repeat: matches 120. .308 of consensus"
6588. .6889
/feature="L1MC/D repeat: matches 5432. .5528 of consensus"
/feature="AlusX repeat: matches 1. .303 of consensus"
Join(7423. .7784,9960. .10039,10306. .10394,10917. .11329,
34968. .35059,36569. .36724,37531. .39351)
/feature="KIAA0964"
Join(<7423. .7784,9960. .10039,10917. .11329,34968. .35059,
36569. .36724,37531. .39351)
/feature="KIAA0964"
/product="dJ977B1.6.2 (KIAA0964 (ortholog of rat
PSD-95/SAP90-associated protein 4))"
/feature="match: cDNAs: Em:AH023181"
/feature="match: cDNAs: Em:AH023181"
/evidence="not experimental"
Join(<7423. .7784,10306. .10394,10917. .11329,34968. .35059,
36569. .36724,37531. .39351)
/feature="KIAA0964"
/product="dJ977B1.6.1 (KIAA0964 (ortholog of rat
PSD-95/SAP90-associated protein 4))"
/feature="match: cDNAs: Em:U67140 Em:AB000277 Em:AB000276
Em:AB000275 Em:AB023181 Em:AF008204 Em:AB003594 Em:U67988
Em:U67137 Em:U67138 Em:U67987 Em:U67139 Em:AF131778
match: ESTs: Em:AI811867 Em:AI620816 Em:AA496750
Em:AA647532 Em:AI747584 Em:AA931146 Em:AI887761
Em:AA154844 Em:AA278057 Em:AI202736 Em:AI497978
Em:AA191427 Em:AA611645 Em:AA760107 Em:AA959160
Em:AA446940 Em:AI190916 Em:U07912 Em:AA558202 Em:AI748854
Em:AI162933 Em:AI840413 Em:AI347137 Em:AA571014
```

```

Em:AA870776 Em:AA144662 Em:AA126125 Em:AA496805 Em:W62099
Em:AA443364 Em:AA874407 Em:AI1986170 Em:AI840039
Em:AA943725 Em:AA955674 Em:AI143631 Em:AA760121
Em:AA050197 Em:R4650 Em:R135758 Em:R4650 Em:R4650 Em:AA433898
Em:R4251 Em:AI918247 Em:R135758 Em:R4650 Em:R4650 Em:R4650
Em:AA824392 Em:R13930 Em:AI948469 Em:AA861813 Em:AA602005
Em:AI201639 Em:W45187 Em:W45187 Em:W45187 Em:W45187 Em:W45187
Em:R12936 Em:R36776 Em:R12916 Em:R12916 Em:R12916 Em:R12916
Em:AA126063 Em:AA434575 Em:AI916248 Em:W75144 Em:W07614
Em:AI972603 Em:AA434575 Em:AI1500286 Em:AI201649
Em:AI143454 Em:AI128723 Em:AI179528 Em:AA031035 Em:W08824
Em:R52222 Em:R51856 Em:R56064 Em:W42146 Em:W49317
Em:R6101 Em:R41361 Em:AA064703 Em:AA295898 Em:AL049119
Em:AA225257 Em:AA046329 Em:AI652021 Em:AA082413 Em:R05501
Em:AA4297 Em:AA046329 Em:AI652021 Em:AA082413 Em:R05501
Em:AI017860 Em:W58510 Em:W58510 Em:W58510 Em:W58510 Em:W58510
/evidence="not experimental"
Join(<7423. .7784,9960. .10039,10917. .11329,34968. .35059,
36569. .36724,37531. .37749)
/feature="KIAA0964"
/feature="isoform 2"
continues in dJ996C2 (AL445705) and BK2182L9 (AL390374)
match: proteins: Tr:Q9Y2H0"
/codon_start=3
/evidence="not experimental"
/product="dJ977B1.6.2 (KIAA0964 (ortholog of rat
PSD-95/SAP90-associated protein 4))"
/db_xref="GI:1132333"
/db_xref="GOA:Q9Y2H0"
/db_xref="GOA:Q9Y2H0"
/translation="SSCIVAYKTPPPVPPRTTSTKPFISVTVQSTSAQTYLDSQD
HKSEVTSQSLNSDSDLSSTPSPVTRGVGVAPEAPPPHAAKSKSQGLTSS
ESHPEAPKRLSSIGIQERTRRGSHLSDNPKAIDVWAPSSSVPSHMSRRD
TSDQDANDSSKSSERSLPDCTPHNPSISIDAGPQAPKIAQIKENLSYGDNSDDA
LEASLPDPWLETSSSPAPQACRDGWFLLKLOAETRLGWCCOMDKET
KENNLSSEVLGKVLAVGSAOLLSQKQFQGLCEONLNDANPRTAQDLGFWDI
LQSLIEDISMKFDEHLKANSWQLVETPEKKEKPPPPVPPKPAKSPKAVSRDKA
SDASDKQKQKRLAAKRAASVRQNSATSDSIIEIYVPEAQL"
Join(<7423. .7784,10306. .10394,10917. .11329,34968. .35059,
36569. .36724,37531. .37749)
/feature="KIAA0964"
/feature="isoform 1"
continues in dJ996C2 (AL445705) and BK2182L9 (AL390374)
match: proteins: Tr:O14488 Tr:O14564 Tr:O14489 Tr:P97839
Tr:O14430 Tr:P97836 Tr:P97837 Tr:O54773 Tr:P97841
Tr:P97835"
/codon_start=3
/evidence="not experimental"
/product="dJ977B1.6.1 (KIAA0964 (ortholog of rat
PSD-95/SAP90-associated protein 4))"
/db_xref="GI:6983730"
/db_xref="GI:6983730"
/db_xref="GOA:Q9Y2H0"
/db_xref="GOA:Q9Y2H0"
/translation="SSCIVAYKTPPPVPPRTTSTKPFISVTVQSTSAQTYLDSQD
HKSEVTSQSLNSDSDLSSTPSPVTRGVGVAPEAPPPHAAKSKSQGLTSS
ESHPEAPKRLSSIGIQERTRRGSHLSDNPKAIDVWAPSSSVPSHMSRRD
TSDQDANDSSKSSERSLPDCTPHNPSISIDAGPQAPKIAQIKENLSYGDNS
DALLASLPPDPWLETSSSPAPQACRDGWFLLKLOAETRLGWCCOMDKET
KENNLSSEVLGKVLAVGSAOLLSQKQFQGLCEONLNDANPRTAQDLGFWDI
LQSLIEDISMKFDEHLKANSWQLVETPEKKEKPPPPVPPKPAKSPKAVSR
DKASDASDKQKQKRLAAKRAASVRQNSATSDSIIEIYVPEAQL"
1065. .10112
/feature="6 copies 8 mer cgcctgc 79% conserved"
11537. .11753
/feature="MIR repeat: matches 11. .236 of consensus"
11767. .11782
/feature="L2 repeat: matches 2692. .2708 of consensus"
11783. .12083
/feature="AlusX repeat: matches 1. .300 of consensus"
12084. .12121
/feature="L2 repeat: matches 2661. .2692 of consensus"
```

```

repeat_region 12631..12714
/notes="HALL repeat: matches 1589..1565 of consensus"
repeat_region 13077..13367 repeat: matches 3..290 of consensus"
/notes="Aludo repeat: matches 1..304 of consensus"
repeat_region 13440..13742
/notes="AlusX repeat: matches 1..304 of consensus"
repeat_region 13932..14197
/notes="lMB7 repeat: matches 5915..5173 of consensus"
repeat_region 14198..14512
/notes="AluY repeat: matches 1..307 of consensus"
repeat_region 14513..14756
/notes="lMB3 repeat: matches 5650..5915 of consensus"
repeat_region 14789..15705
/notes="lMB4 repeat: matches 5162..5185 of consensus"
repeat_region 15706..16004

Query Match 12.8%; Score 152; DB 9; Length 145068;
Best Local Similarity 100.0%; Pred. No. 3e-76;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 778 AGGCTCTTACTCTCTGTCAGTCGGCTCAGCGCCCTGCATCTCGGACCGGATCAGACA 837
Db 126094 AGGCTCTTACTCTCTGTCAGTCGGCTCAGCGCCCTGCATCTCGGACCGGATCAGACA 126095

Qy 938 CTACAGGATCCACTGCTTGAACAATGGCTGTGATCTACCGCGGCTTCACTTCCC 897
Db 126034 CTACAGGATCCACTGCTTGAACAATGGCTGTGATCTACCGCGGCTTCACTTCCC 125975

Qy 898 CTCACCTCCAGGCGCCCTGGTGACCATTAATCTCTG 929
Db 125974 CTCACCTCCAGGCGCCCTGGTGACCATTAATCTCTG 125943

```

Search completed: February 20, 2004, 07:24:47  
Job time : 4588 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 01:41:55 ; Search time 503 Seconds  
(without alignments)  
9991.289 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccaggacc.....tctcttgatgatgctag 1183

Scoring table: OLIGO NUC  
Gapop 50.0 , Gapext 60.0

Searched: 337863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002as.\*
- 7: Geneseq2003as.\*
- 8: Geneseq2003bs.\*
- 9: Geneseq2003cs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183	100.0	1183	6	ABK61465 Human CDN
2	1132	95.7	2567	6	AAL43980 Human Src
3	735	62.1	786	6	AAL44089 Human mod
4	735	62.1	786	6	ABQ74343 Human Src
5	657	55.5	763	6	ABQ98670 Human ORF
6	592	50.0	837	3	AAC77202 Human ORF
7	483	40.8	737	6	AAL44090 Mouse MAR
8	468	39.6	1413	6	ABQ99374 Human cod
9	405	34.2	603	5	AS74748 DNA encod
10	401	33.9	2049	5	AS74750 DNA encod
11	348	23.4	444	6	ABQ98669 Human ORF
12	341	28.8	875	6	ABQ99151 Human ORF
13	244	20.6	445	5	AS74747 DNA encod
14	141	11.9	211	5	AS70181 DNA encod
15	134	11.3	432	4	AAI12879 Probe #28
16	134	11.3	432	4	ABA54580 Human foe
17	134	11.3	432	4	AAI34236 Probe #29
18	134	11.3	432	4	ABA4128 Human bre
19	134	11.3	432	4	ABA24363 Probe #28
20	134	11.3	432	4	AAL28314 Human bon
21	134	11.3	432	4	AAK02872 Human bra
22	134	11.3	432	4	ABS27912 Human liv
23	134	11.3	432	5	AAI02797 Probe #27

24	134	11.3	432	6	ABS02823	Human gen
25	134	11.3	448	4	AAI14520	Probe #44
26	134	11.3	448	6	ABS04499	Human gen
27	96	8.1	96	4	AAI22119	Probe #12
28	96	8.1	96	4	ABA67198	Human foe
29	96	8.1	96	4	AAI47414	Probe #16
30	96	8.1	96	4	ABA49284	Human bre
31	96	8.1	96	4	ABA4292	Probe #12
32	96	8.1	96	4	AAK41374	Human bon
33	96	8.1	96	4	AAK15640	Human bra
34	96	8.1	96	4	ABS40966	Human liv
35	96	8.1	96	5	AAI07818	Probe #78
36	96	8.1	96	6	ABS15380	Human gen
37	43	3.6	1348	6	AAL44087	Mouse mod
38	30	2.5	30	6	AAD43983	Human leu
39	26	2.2	26	6	ABK61506	Human NOY
40	25	2.1	25	6	AAL44100	Human mod
41	25	2.1	25	6	AAL44098	Human mod
42	25	2.1	25	6	AAL44099	Human mod
43	23	1.9	23	6	AAL44097	Human mod
44	22	1.9	22	6	ABK61507	Human NOV
45	22	1.9	3070	2	AAK55277	Nucleotid

## ALIGNMENTS

RESULT 1  
ABK61465  
ID : ABK61465 standard; cDNA; 1183 BP.  
XX  
AC ABK61465;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human cDNA encoding protein NOV13.  
XX  
KW Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;  
KW cell signal processing disorder; metabolic pathway modulation disorder;  
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;  
KW uterine cancer; immune response; graft-versus-host disease;  
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;  
KW hyperension; congenital heart defects; multiple sclerosis; inflammation;  
KW Albright hereditary osteodystrophy.  
XX  
OS Homo sapiens.  
XX  
FN WO200216599-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 27-AUG-2001; 2001WO-US026510.  
XX  
PR 25-AUG-2000; 2000US-0228191P.  
PR 08-FEB-2001; 2001US-0267300P.  
PR 20-FEB-2001; 2001US-0269961P.  
PR 20-MAR-2001; 2001US-0277337P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Burgeses CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;  
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;  
XX  
DR WPI; 2002-280937/32.  
XX  
P-PSDB; AAU91308.  
XX  
PT New polypeptides for treating or preventing a disorder associated with  
them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.  
XX  
PS Claim 1; Page 98; 263pp; English.  
XX  
CC The invention relates to an isolated polypeptide (NOVX) a mature form of

CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide  
 CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
 CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it,  
 CC and antibody against it, are useful for treating or preventing (e.g. by  
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOVX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation of Albrecht  
 CC hereditary osteodystrophy and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC for developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence encodes a NOVX protein.

XX  
 SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Query Match 100.0%; Score 1183; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTAGAGCTCCAGAGACCCAGCGCTGTCTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
 DB 1 AGCTAGAGCTCCAGAGACCCAGCGCTGTCTCTGTGACAGAGCTCAAGGGCCCTGGG 60

QY 61 CTTTCCTCCCTGGCTGGCTGTCTGTGGAGGGTTCCCGAGTCAGAAATCCCTAAGGAG 120  
 DB 61 CTTTCCTCCCTGGCTGGCTGTCTGTGGAGGGTTCCCGAGTCAGAAATCCCTAAGGAG 120

QY 121 CATGGGCGAGCTGATCCATCCCTGGTGTACAAACTGCTGACTGCAGACAGATGCTGAGCT 180  
 DB 121 CATGGGCGAGCTGATCCATCCCTGGTGTACAAACTGCTGACTGCAGACAGATGCTGAGCT 180

QY 181 ACCCAACACACCTAGCTCCCTCCCTGAGATCTCCCGAGCTGAGAGAGTTCTGGGTG 240  
 DB 181 ACCCAACACACCTAGCTCCCTCCCTGAGATCTCCCGAGCTGAGAGAGTTCTGGGTG 240

QY 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAAAGGGCCCCCAAGCCCTAACCTGTCCA 300  
 DB 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAAAGGGCCCCCAAGCCCTAACCTGTCCA 300

QY 301 GCCAGAGCTGCTCTCAGCAGAGCTGTCTTCCAGAGCTTTGATGACAAACCAATTTCC 360  
 DB 301 GCCAGAGCTGCTCTCAGCAGAGCTGTCTTCCAGAGCTTTGATGACAAACCAATTTCC 360

QY 361 CTCGATGATGTCTTCTGAGTGTCTCTGCTGAGGAACAATGGGAAGTCTGCCAGCAGAAG 420  
 DB 361 CTCGATGATGTCTTCTGAGTGTCTCTGCTGAGGAACAATGGGAAGTCTGCCAGCAGAAG 420

QY 421 AABAATCTGCGAGCCAGCTGAGTCTCTGCTGAGGAGCCAGGACCTGTGACCAT 480  
 DB 421 AABAATCTGCGAGCCAGCTGAGTCTCTGCTGAGGAGCCAGGACCTGTGACCAT 480

QY 481 GGAAGCAGAGAGAGCAGAGCCAGCGGTGGCCCTGGGAGTTTCCCGGAGGTGGGCC 540  
 DB 481 GGAAGCAGAGAGAGCAGAGCCAGCGGTGGCCCTGGGAGTTTCCCGGAGGTGGGCC 540

QY 541 GGCGAGCTGTGCTGAGATCTGGGAGGACATGACCTGCTCTGAGGATGGAGACTG 600  
 DB 541 GGCGAGCTGTGCTGAGATCTGGGAGGACATGACCTGCTCTGAGGATGGAGACTG 600

QY 601 GTGAGCGGTGCTGTCTGAAGTCTCAGCAGAGAGTATAATCCCGAGCGTCCACGTGGG 660  
 DB 601 GTGAGCGGTGCTGTCTGAAGTCTCAGCAGAGAGTATAATCCCGAGCGTCCACGTGGG 660

QY 661 CAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGAAGTCTGT 720  
 DB 661 CAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGAAGTCTGT 720

QY 721 GTTGTACTCTGGAAACCTTGGAGGGCTTCTCTATCCGGAGAGCCAGACAGAGAGG 780  
 DB 721 GTTGTACTCTGGAAACCTTGGAGGGCTTCTCTATCCGGAGAGCCAGACAGAGAGG 780

QY 781 CTCTTACTCTCTGTCAGTCCGCTCAGCCGCTGATCTCTGGGACCGGATCAGACACTA 840  
 DB 781 CTCTTACTCTCTGTCAGTCCGCTCAGCCGCTGATCTCTGGGACCGGATCAGACACTA 840

QY 841 CAGGATCCACTCCCTTGACATGGCTGTGATCTCAGCGCGCTCACCTTCCCTCCCTC 900  
 DB 841 CAGGATCCACTCCCTTGACATGGCTGTGATCTCAGCGCGCTCACCTTCCCTCCCTC 900

QY 901 ACTCCAGGCCCTGGTGGACCACTTACTCTGAGTGGCGGATGACATCTGCTCCCTACTCAA 960  
 DB 901 ACTCCAGGCCCTGGTGGACCACTTACTCTGAGTGGCGGATGACATCTGCTCCCTACTCAA 960

QY 961 GGAGCCCTCTGTCTGTCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGT 1020  
 DB 961 GGAGCCCTCTGTCTGTCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGT 1020

QY 1021 GACTGTGACAGGACACCACTCACTGGAAAGAGCTGGAGCTGCCCTCTCTGTTTCTGA 1080  
 DB 1021 GACTGTGACAGGACACCACTCACTGGAAAGAGCTGGAGCTGCCCTCTCTGTTTCTGA 1080

QY 1081 AGTGTGCCAGAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTA 1140  
 DB 1081 AGTGTGCCAGAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTA 1140

QY 1141 CATCAGCTGAATGACAGGGTGTCTCTTTGGATGATGCTCTAG 1183  
 DB 1141 CATCAGCTGAATGACAGGGTGTCTCTTTGGATGATGCTCTAG 1183

## RESULT 2

AAD43980  
 ID AAD43980 standard; cDNA; 2567 BP.

XX AAD43980;

DT 13-DEC-2002 (first entry)

XX Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.

Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 multiple sclerosis; asthma; acute respiratory distress syndrome;  
 pulmonary disorder; dermatological; neuroprotective; gene; ss.

XX Homo sapiens.

PH Key Location/Qualifiers  
 CDS 415..1200

FT /\*tag= a

FT /product= "Human SLAP-2"

XX WO200242457-A1.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US043367.

XX 22-NOV-2000; 2000US-0252545P.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;

XX WPI: 2002-463632/49.  
 DR P-PSDB; AAE26357.  
 DR  
 XX Novel substantially purified human SH2/SH3-domain-containing adapter  
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic  
 PT intervention in immunological and inflammatory disorders and cancer.  
 XX  
 PS Claim 2; Fig 1; 85pp; English.  
 XX  
 CC The invention relates to a substantially purified human SH2/SH3-domain-  
 CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-  
 CC 2). The invention is useful for treating an immune disorder involving  
 CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is  
 CC useful for screening for antagonists or inhibitors of the interaction of  
 CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or  
 CC preventing diseases or disorders associated with aberrant or uncontrolled  
 CC cellular signal transduction, for determining those cellular signalling  
 CC molecules which associate with hSLAP-2 and which provide critical signals  
 CC for cell activation, and as effectors in methods to affect T- cell  
 CC activation. The invention is useful in screening assays to identify and  
 CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for  
 CC potential use to treat autoimmune diseases which may be caused by  
 CC hyperactivated B cells, as well as to treat diseases which may be caused  
 CC by hyperactivated T cells, in addition to other immune system related  
 CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,  
 CC inflammation disorders, diseases and conditions, rheumatoid arthritis,  
 CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's  
 CC and ulcerative colitis), allergies, particularly those involving  
 CC hyperactivity of B-cells and T-cells, or other immune cells, such as  
 CC mast cells or eosinophils, autoimmune diseases such as systemic lupus  
 CC erythematosus and multiple sclerosis, pulmonary diseases including  
 CC asthma, acute respiratory distress syndrome, and chronic obstructive  
 CC pulmonary disorder, tissue/organ rejection and cancer. The invention is  
 CC useful in gene therapy. The present sequence is human SLAP-2 cDNA  
 XX  
 SQ Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;  
 Query Match 95.7%; Score 1132; DB 6; Length 2567;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 1 AGCTAGAGCTCCAGGACCCCAAGCCCTGCTGCTGTGACAGAGCTCAAGAGCCCTGGG 60  
 18 AGCTAGAGCTCCAGGACCCCAAGCCCTGCTGCTGTGACAGAGCTCAAGAGCCCTGGG 77  
 61 CCTTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 78 CCTTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 137  
 121 CATGGGGAGCTGATCCATCCCTGCTGTGTACAACTGTGTGCTGCTGCTGCTGCTGCTGCT 180  
 138 CATGGGGAGCTGATCCATCCCTGCTGTGTACAACTGTGTGCTGCTGCTGCTGCTGCTGCT 197  
 181 ACCCAAAACCAACACCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 198 ACCCAAAACCAACACCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257  
 241 TCTTAGGACCAAGGACACTGGGAGACTTCCAGAGGGCCGCCAAAGCCCTAACCTGTCCA 300  
 258 TCTTAGGACCAAGGACACTGGGAGACTTCCAGAGGGCCGCCAAAGCCCTAACCTGTCCA 317  
 301 GCCAGAGATGGCTCTGACAGAGAGCTGCTTCCCAAGCCCTTGTATGACAAACCAATTTCC 360  
 318 GCCAGAGATGGCTCTGACAGAGAGCTGCTTCCCAAGCCCTTGTATGACAAACCAATTTCC 377  
 361 CTCGATGATGCTTCTGAGTGCTCTGCTGAGGACCAATGGGAAGCTGCCAGCAGAGAG 420  
 378 CTCGATGATGCTTCTGAGTGCTCTGCTGAGGACCAATGGGAAGCTGCCAGCAGAGAG 437  
 421 AAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTCTCCAAAGGCCAGGAGCCTGTGACCAT 480  
 438 AAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTCTCCAAAGGCCAGGAGCCTGTGACCAT 497

QY 481 GGAAGCAG 540  
 DB 498 GGAAGCAG 557  
 QY 541 GGCCGAGAGCTGTGCTGAG 600  
 DB 558 GGCCGAGAGCTGTGCTGAG 617  
 QY 601 GTGAG 660  
 DB 618 GTGAG 677  
 QY 661 CAAAGTCTCCCAAG 720  
 DB 678 CAAAGTCTCCCAAG 737  
 QY 721 GTTGTACTTGGGAG 780  
 DB 738 GTTGTACTTGGGAG 797  
 QY 781 CTCTTACTTCTGTGAG 840  
 DB 798 CTCTTACTTCTGTGAG 857  
 QY 841 CAGGATCCCAAG 900  
 DB 858 CAGGATCCCAAG 917  
 QY 901 ACTCCAG 960  
 DB 918 ACTCCAG 977  
 QY 961 GGAG 1020  
 DB 978 GGAG 1037  
 QY 1021 GACTGTGAG 1080  
 DB 1038 GACTGTGAG 1097  
 QY 1081 AGTGTGAG 1140  
 DB 1098 AGTGTGAG 1157  
 QY 1141 CATCAGCCTGAATGAG 1183  
 DB 1158 CATCAGCCTGAATGAG 1200  
 RESULT 3  
 AAL44089  
 ID AAL44089 standard; cDNA; 786 BP.  
 XX AAL44089;  
 AC AAL44089;  
 XX  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Human modulator of antigen receptor signalling protein coding sequence.  
 XX  
 KW Human; gene; ss; gene therapy; modulator of antigen receptor signalling;  
 KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;  
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 OS Homo sapiens.  
 XX  
 FR Key Location/Qualifiers  
 FT CDS 1..786  
 FT /\*tag= a  
 FT /product= "Human MARS protein"  
 XX  
 PN WO200242452-A2.

XX PD 30-MAY-2002.  
 XX PF 26-NOV-2001; 2001WO-CA001662.  
 XX PR 27-NOV-2000; 2000CA-02324663.  
 XX PA (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 XX PI Mcglade JC, Loreto MP;  
 XX DR WPI: 2002-566564/60.  
 XX DR P-PSDB; A0015457.  
 XX PT New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX PS Claim 12; Page 75; 110pp; English.  
 XX CC The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present cDNA sequence encodes a human MARS protein  
 XX  
 XX SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;  
 Query Match 62.1%; Score 735; DB 6; Length 786;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 398 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCAAGCCCAAGCTTGATTCCTCTGTC 457  
 DB 1 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCAAGCCCAAGCTTGATTCCTCTGTC 60  
 QY 458 CAAGCCAGGACCTGTGACCATGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 517  
 DB 61 CAAGCCAGGACCTGTGACCATGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAT 120  
 QY 518 GGCAGTTTCCGCGCAGAGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCATTGACC 577  
 DB 121 GGCAGTTTCCGCGCAGAGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCATTGACC 180  
 QY 578 ATCGTCTCTGAGATGGAGAGTGTGGACGCTGTCTGAGTCTCAGCAGAGAGAT 637  
 DB 181 ATCGTCTCTGAGATGGAGAGTGTGGACGCTGTCTGAGTCTCAGCAGAGAGAT 240  
 QY 638 AACATCCCGCAGGCTCCAGTGGGGAAGTCTCCATGGGTGGCTGTATGAGGCGCTGAGC 697  
 DB 241 AACATCCCGCAGGCTCCAGTGGGGAAGTCTCCATGGGTGGCTGTATGAGGCGCTGAGC 300  
 QY 698 AGGGAGAGCAGAGAGAACTCTGTGTGTACTGGGAACCTCGAGGGGCTTCTCTCATC 757  
 DB 301 AGGGAGAGCAGAGAGAACTCTGTGTGTACTGGGAACCTCGAGGGGCTTCTCTCATC 360  
 QY 758 CGGGAGAGCAGAGAGAGTCTTACTCTGTCTGTCAGTCCGCTCAGCCGCTTGA 817  
 DB 361 CGGGAGAGCAGAGAGAGTCTTACTCTGTCTGTCAGTCCGCTCAGCCGCTTGA 420  
 QY 818 TCTGGGACCGGATCAGACACTACAGATCCACTGCGCTTGACATGGCTGGCTGTACATC 877  
 DB 421 TCTGGGACCGGATCAGACACTACAGATCCACTGCGCTTGACATGGCTGGCTGTACATC 480  
 QY 878 TCACGCGGCTCAGCTTCCCTCTACTCCAGCCCTGTGTGACCATTTACTGAGCTGGG 937  
 DB 481 TCACGCGGCTCAGCTTCCCTCTACTCCAGCCCTGTGTGACCATTTACTGAGCTGGG 540  
 QY 938 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGCTCTGAGAGGGGCTGCCGCTCCCT 997

DB 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTCTCTGAGAGGCTGCGCCGCTCCCT 600  
 QY 998 GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACACTCACTGGAAGAGCTG 1057  
 DB 601 GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACACTCACTGGAAGAGCTG 660  
 QY 1058 GACAGTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTTCTCAGTGGGT 1117  
 DB 661 GACAGTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTTCTCAGTGGGT 720  
 QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATACGAGGCTGCTCTTTGGATGAT 1177  
 DB 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATACGAGGCTGCTCTTTGGATGAT 780  
 QY 1178 GCTTAG 1183  
 DB 781 GCTTAG 786  
 RESULT 4  
 ABQ74343  
 ID ABQ74343 standard; cDNA; 786 BP.  
 XX AC ABQ74343;  
 XX DT 15-OCT-2002 (first entry)  
 XX DE Human Src-like inhibitory molecule (SLIM) encoding cDNA.  
 XX KW Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;  
 KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;  
 KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;  
 KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
 KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;  
 KW gene; ss.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 FT CDS 1..786  
 FT /tag= a "SLIM"  
 FT /product= "SLIM"  
 FT /note= "Src-like inhibitory molecule"  
 XX PN W020255707-A2.  
 XX PD 18-JUL-2002.  
 XX PF 10-JAN-2002; 2002WO-US000718.  
 XX PR 10-JAN-2001; 2001US-0260953P.  
 XX PA (RIGE-) RIGEL PHARM INC.  
 XX PI Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
 XX WPI: 2002-575432/61.  
 XX P-PSDB; ABP52187.  
 XX PT New src-like inhibitory molecule protein, useful for treating  
 PT immunodeficiency disorders and inflammatory disorders, comprises N-  
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.  
 XX PS Claim 3; Fig 2A; 91pp; English.  
 XX CC The present sequence encodes the human Src-like inhibitory molecule  
 CC (SLIM) protein (I). The present invention describes a SLIM protein  
 CC comprising an N-terminal myristylation sequence, an N-terminal SH2  
 CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising  
 CC an N-terminal myristylation sequence and an N-terminal SH2 domain which  
 CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and



anti-HIV activities, and can be used as a modulator of lymphocyte activation, and of upregulation of a Cbl target protein, and in gene therapy. (I) is useful for screening a bioactive agent capable of binding to SHM. (I) is also useful for screening a bioactive agent capable of modulating SHM binding. (I) or its fragments is useful in the study or in the treatment of conditions which involves this function or dysregulation of SHM protein activity, i.e. to diagnose, treat or prevent SHM associated disorders. (I) or the polynucleotide encoding it (II) is useful for modulating leukocyte and/or platelet activation, for modulating antigen receptor-induced signalling and activation in leukocyte and/or platelets and for modulating antigen receptor-induced signalling and activation in lymphocytes and/or mast cells. (I) or (II) is also useful for modulating the basal activity of lymphocytes. (I) or (II) is useful in the treatment of immunodeficiency disorders, such as acquired immunodeficiency syndrome (AIDS), for the prevention and treatment of acute inflammatory disorders, chronic inflammatory disorders, autoimmune disorder and transplant rejection

Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other:  
XX  
SQ

Db 721 CTCGCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACAGGCTGTCCTTTGGATGAT 780

QY 1178 GCCTAG 1183

Db 781 GCCTAG 786

RESULT 5  
ABQ98670  
ID ASQ98670 standard; DNA; 763 BP.  
XX  
XX  
XX  
XX  
DT 04-NOV-2002 (first entry)  
XX  
DE Human QR477 coding sequence.

XX	Homo sapiens.
CS	
XX	US2002082206-A1.
PN	
XX	27-JUN-2002.
XX	
XX	30-MAY-2001; 2001US-00867550.
PF	
XX	
PR	30-MAY-2000; 2000US-0208427P.

xx New polypeptide designated ORFX are present in human atherogenic cells  
PT and are useful to prevent and treat ORFX-associated disorders including  
PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
PT inflammatory disease.

PS Claim 2; SEQ ID NO 953; 78pp; English.

The present invention relates to novel human ORFX polypeptides and their coding sequences (ABP631-ABP64681 and ABQ98194-ABQ99267). The sequences were discovered in human atherosclerotic cells, in particular in platelets and human umbilical vein endothelial cells (HUVEC) and are expressed in many other tissues as well. Atherosclerotic cells are cells which have the potential to develop atherosclerotic plaques. The ORFX polypeptides and nucleic acids are useful for treating or preventing a pathological condition associated with an ORFX-associated disorder, e.g. cancer, cardiovascular disease, allergy, autoimmune disease, wound healing, blood coagulation disorders or inflammatory disorders. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at [seqdata.uspto.gov/sequence.html?docID=20020082205](http://seqdata.uspto.gov/sequence.html?docID=20020082205)

Sequence 763 BP: 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 55.5%; Score 657; DB 6; Length 763;  
Best Local Similarity 99.7%; Pred. No. 5.6e-300;  
Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0

OV 117 GGAGCATGGGGAGCTGATCCATCCCTGGTGTACAAACTGCTCACTGCAGACAGATGCTG 176

Db 5 GGAGCATGGGCGAGTGCATCCCTGGTGTACAACTGCTGCTGACGACAGATGCTG 64  
 QY 177 AGCTACCCCAACACCACTAGCTCTCCCTGAGATCTCCAGGCTGAGAGATTCTG 236  
 Db 65 AGCTACCCCAACACCACTAGCTCTCCCTGAGATCTCCAGGCTGAGAGATTCTG 124  
 QY 237 GGTGTCCTAGACCAAGGACACTGGCAGACTTCCAGAAAGGGCCCCAAAGCCCTAACCTG 296  
 Db 125 GATGTCTTAGACCAAGGACACTGGCAGACTTCCAGAAAGGGCCCCAAAGCCCTAACCTG 184  
 QY 297 TCCAGCCAGACGATGCTGCTCAGCAGAGCTGTCTTCCAGAGCTTGTGATGACAAACCAAT 356  
 Db 185 TCCAGCCAGACGATGCTGCTCAGCAGAGCTGTCTTCCAGAGCTTGTGATGACAAACCAAT 244  
 QY 357 TTCCCTCGATGATGCTTCTGAGTCTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCA 416  
 Db 245 TTCCCTCGATGATGCTTCTGAGTCTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCA 304  
 QY 417 CAAGAAATCTCTCCAGCCCAAGCTTGTGATGCTTCCAGAGCTTGTGATGACAAACCAAT 476  
 Db 305 CAAGAAATCTCTCCAGCCCAAGCTTGTGATGCTTCCAGAGCTTGTGATGACAAACCAAT 364  
 QY 477 CCATGGAGAGCAGAGAGAGCAAGCCACAGCCCTGGCCCTGGCCAGTCTTCCCGCAGGTG 536  
 Db 365 CCATGGAGAGCAGAGAGAGCAAGCCACAGCCCTGGCCCTGGCCAGTCTTCCCGCAGGTG 424  
 QY 537 GCCCGGCCGAGCTGCTGAGACTCGGGAGGCAATGACCATGCTGCTGAGGATGGAG 596  
 Db 425 GCCCGGCCGAGCTGCTGAGACTCGGGAGGCAATGACCATGCTGCTGAGGATGGAG 484  
 QY 597 ACTGGTGAAGCGTCTGCTGAGTCTCAGGAGAGAGATATACATCCCGCAGGTCCACG 656  
 Db 485 ACTGGTGAAGCGTCTGCTGAGTCTCAGGAGAGAGATATACATCCCGCAGGTCCACG 544  
 QY 657 TGGGCAAAAGTCTCCATGGGTGGTGTATGAGGCTGAGGAGGAGAAAGCAGAGGAAC 716  
 Db 545 TGGGCAAAAGTCTCCATGGGTGGTGTATGAGGCTGAGGAGGAGAAAGCAGAGGAAC 604  
 QY 717 TGCTGTGTTTACCTGGGAACTTGGAGGCGCTTCTCTATCGGAGAGAGCCAGACAGGA 776  
 Db 605 TGCTGTGTTTACCTGGGAACTTGGAGGCGCTTCTCTATCGGAGAGAGCCAGACAGGA 664  
 QY 777 GAGGCTCTTACTCTCTAGTCCGCTCAGCGCCCTGCATCTCGGAGCCGATCAGAC 836  
 Db 665 GAGGCTCTTACTCTCTAGTCCGCTCAGCGCCCTGCATCTCGGAGCCGATCAGAC 724  
 QY 837 ACTACAGATCCACTGCTTGACAAATGGCTGGCTGTACA 875  
 Db 725 ACTACAGATCCACTGCTTGACAAATGGCTGGCTGTACA 763

## RESULT 6

AAC77202

ID AAC77202 standard; cDNA; 837 BP.

XX AC

AC AAC77202;

XX DT

DE 08-FEB-2001 (first entry)

Human OREF ORF2757 polynucleotide sequence SEQ ID NO:5513.  
 Human; open reading frame; OREF; detection; cytosolic; hepatotropic;

vulnary; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;  
 anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;  
 immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;  
 hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 cholesterol ester storage; systemic lupus erythematosus; infection;  
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 thrombosis; contraceptive; ss.

Homo sapiens.

WO200058473-A2.

FN XX

PD XX

PF XX

31-MAR-2000; 2000WO-US008621.

31-MAR-1999; 99US-0127607P.

02-APR-1999; 99US-0127636P.

05-APR-1999; 99US-0127728P.

30-MAR-2000; 2000US-00540763.

(CURA-) CURAGEN CORP.

Shimkets RA, Leach M;

WPI; 2000-602362/57.

P-PSDB; AAB42993.

XX XX

PT PT

PT PT

XX XX

PS XX

XX XX

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

XX XX

SQ

Sequence 837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;

Query Match 50.0%; Score 592; DB 3; Length 837;

Best Local Similarity 99.6%; Pred. No. 2.8e-269;

Matches 742; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 439 AAGCTTGAGTCTCTCTGTCACAGGCGGAGGAGCTGTGACCATGGAAGCAGAGAGACAA 498

Db 3 AAGCTTGAGTCTCTCTGTCACAGGCGGAGGAGCTGTGACCATGGAAGCAGAGAGACAA 62

QY 499 GGCCACAGCCGCGGCGCTGGGAGTTCCTCCGAGGTTGGCCGCGGAGCTGCTGCTGAG 558

Db 63 GGCCACAGCCGCGGCGCTGGGAGTTCCTCCGAGGTTGGCCGCGGAGCTGCTGCTGAG 122

QY 559 ACTCGGGAGCCATTGACCATCTGCTCTGAGGATGGAGACTGTGGACGGTGTGCTGTA 618

Db 123 ACTCGGGAGCCATTGACCATCTGCTCTGAGGATGGAGACTGTGGACGGTGTGCTGTA 182

QY 619 AGTCTCAGCGAGAGATATACATCCCGGAGCTCCAGTGGGCAAGTCTCCCATGGGTG 678

Db 183 AGTCTCAGCGAGAGATATACATCCCGGAGCTCCAGTGGGCAAGTCTCCCATGGGTG 242

QY 679 GCTGTATGAGGCTCTGAGCAGGAGAAAGCAGAGAACTGCTGTGTGTACCTGGAAACC 738  
 Db 243 GCTGTATGAGGCTCTGAGCAGGAGAAAGCAGAGAACTGCTGTGTGTACCTGGAAACC 302  
 QY 739 TGGAGGGGCTTCTCTCATCCGAGAGCCAGACAGGAGGCTTACTCTCTGTCACT 798  
 Db 303 TGGAGGGGCTTCTCTCATCCGAGAGCCAGACAGGAGGCTTACTCTCTGTCACT 362  
 QY 799 CCCTCTCAGCGGCTCTGATCTCTGGACCGGATCAGACACTACAGATCCACTGCCCTGA 858  
 Db 363 CCCTCTCAGCGGCTCTGATCTCTGGACCGGATCAGACACTACAGATCCACTGCCCTGA 422  
 QY 859 CAATGCTGGCTGTACATCTCAACGCGGCTCACTTCCCTCACTCCAGGCGCTGTGGGA 918  
 Db 423 CAATGCTGGCTGTACATCTCAACGCGGCTCACTTCCCTCACTCCAGGCGCTGTGGGA 482  
 QY 919 CCAATGCTGGCTGTACATCTCAACGCGGCTCACTTCCCTCACTCCAGGCGCTGTGGGA 978  
 Db 483 CCAATGCTGGCTGTACATCTCAACGCGGCTCACTTCCCTCACTCCAGGCGCTGTGGGA 542  
 QY 979 GAGGGCTGGCGGCTCTCCCTGGCAAGGATATACCCCTACTGTGACTGTGACAGGACACC 1038  
 Db 543 GAGGGCTGGCGGCTCTCCCTGGCAAGGATATACCCCTACTGTGACTGTGACAGGACACC 602  
 QY 1039 ACTCACTGGAAGCTGTGACAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1098  
 Db 603 ACTCACTGGAAGCTGTGACAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 662  
 QY 1099 GTCTCTTCTCAGTGAGGCTCTCCGAGAGTCCCTCACTCTCTCTCTCTCTCTCTCTCTCT 1158  
 Db 663 GTCTCTTCTCAGTGAGGCTCTCCGAGAGTCCCTCACTCTCTCTCTCTCTCTCTCTCTCT 722  
 QY 1159 GGCTGTCTCTTTGGATGAGCTAG 1183  
 Db 723 GGCTGTCTCTTTGGATGAGCTAG 747

RESULT 7  
 AAL44090  
 ID AAL44090 standard; cDNA; 737 BP.  
 XX  
 AC AAL44090;  
 XT  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Mouse MARS short isoform protein coding sequence.  
 KW Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;  
 KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;  
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..633  
 FT /\*tag= a  
 FT /product= "mouse MARS short isoform protein"  
 XX  
 DN WO200242452-A2.  
 XX  
 PD 30-MAY-2002.  
 XX  
 PF 26-NOV-2001; 2001WO-CA001662.  
 XX  
 PR 27-NOV-2000; 2000CA-02324663.  
 XX  
 PA (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 XX  
 FI Mcglade JC, Loreto MP;  
 XX  
 DR WPI; 2002-566564/60.  
 DR P-PSDB; AA015458.

XX New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX  
 PS Claim 9; Page 77; 110pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present cDNA sequence encodes a mouse MARS protein  
 XX  
 SQ Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;  
 Query Match 40.8%; Score 483; DB 6; Length 737;  
 Best Local Similarity 99.8%; Pred. No. 8.9e-218;  
 Matches 533; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGTCTCTGTCT 457  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGTCTCTGTCT 60  
 QY 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGTGGCCCTG 517  
 Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGTGGCCCTG 120  
 QY 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 577  
 Db 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180  
 QY 578 ATCTGTCTCTGAGGATGGAGACTGGTGACGGTGTGTCTGAACTCTCAGGCAGAGAGTAT 637  
 Db 181 ATCTGTCTCTGAGGATGGAGACTGGTGACGGTGTGTCTGAACTCTCAGGCAGAGAGTAT 240  
 QY 638 AACATCCCCCAGCTCCAGCTGGGCAAGTCTCCATGGGTGGTGTATGAGGCCCTGAGC 697  
 Db 241 AACATCCCCCAGCTCCAGCTGGGCAAGTCTCCATGGGTGGTGTATGAGGCCCTGAGC 300  
 QY 698 AGGAGAAAGCAGAGGAACTGCTGTGTACCTGGGAACCCCTGGAGGGCCCTTCTCTCATC 757  
 Db 301 AGGAGAAAGCAGAGGAACTGCTGTGTACCTGGGAACCCCTGGAGGGCCCTTCTCTCATC 360  
 QY 758 CGGAGAGCCAGACCCAGGAGGCTCTTACTCTCTCTCAGTCCGCTCAGCGCCCTGCA 817  
 Db 361 CGGAGAGCCAGACCCAGGAGGCTCTTACTCTCTCTCAGTCCGCTCAGCGCCCTGCA 420  
 QY 818 TCCTGGGACCGGATCAGACACTCAGGATCCACTGCTTGACAATGGCTGGTGTACATC 877  
 Db 421 TCCTGGGACCGGATCAGACACTCAGGATCCACTGCTTGACAATGGCTGGTGTACATC 480  
 QY 878 TCACCGGCTCTCAGCTTCCCTCAGTCCAGCCCTCGTGGAGCCATTACTCTGAG 931  
 Db 481 TCACCGGCTCTCAGCTTCCCTCAGTCCAGCCCTCGTGGAGCCATTACTCTGAG 534  
 RESULT 8  
 ABC99374  
 ID ABC99374 standard; cDNA; 1413 BP.  
 XX  
 AC ABC99374;  
 XX  
 DT 25-FEB-2003 (first entry)  
 XX  
 DE Human coding sequence SEQ ID 107.  
 XX  
 KW Human; expressed sequence tag; EST; chromosome 20;  
 KW haematopoietic disorder; central nervous system disease; viral infection;  
 KW peripheral nervous system disease; non-healing wound; infectious disease;

immune deficiency; immune disorder; bacterial infection; allergy; cancer; fungal infection; autoimmune disorder; coagulation disorder; neutropenic; anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective; cytoskeletal; haemostatic; virulence; antibacterial; fungicide; immunostimulant; cerebroprotective; gene therapy; gene, ss.

OS Homo sapiens.

XX WO200259250-A2.

XX 01-AUG-2002.

XX 16-NOV-2001; 2001WO-US042950.

XX 17-NOV-2000; 2000US-00714936.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA; Ren P, Xue AJ, Yang Y, Wehrman T, Drmanac RT; WPI; 2002-590824/63.

XX N-PSDB; ABP64788.

XX New isolated polynucleotide, useful in research, diagnostic or therapeutic methods, e.g. preventing or treating disorders involving aberrant protein expression or biological activity.

XX Claim 1; SEQ ID NO 107; 394pp; English.

The present invention relates to novel human coding sequences (ABQ99268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in therapeutic, diagnostic and research methods. The polynucleotides may be used in the field of molecular biology as hybridisation probes, primers for PCR, for chromosome and gene mapping, for the recombinant production of protein, or in generation of anti-sense DNA or RNA. The polynucleotides are useful in diagnostics as expressed sequence tags (ESTs) for identifying expressed genes or for physical mapping of the human genome. The proteins may be used as molecular weight markers, or as nutritional sources or supplements. The proteins may be used to maintain and expand cell population in a totipotent or pluripotent state useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals or the development of bio-sensors. The polynucleotides and proteins are useful for preventing, treating or ameliorating disorders involving aberrant protein expression or biological activity, e.g. haematopoietic disorders, central/peripheral nervous system diseases, mechanical and traumatic disorders, non-healing wounds, immune deficiencies and disorders, infectious diseases caused by viral, bacterial or fungal infection, autoimmune disorders, allergic reactions and conditions, coagulation disorders, or cancer. The polynucleotide sequences of the invention were assembled from ESTs isolated mainly by sequencing by hybridisation, and in some cases, sequences obtained from one or more public databases. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;

Query Match 39.6%; Score 468; DB 6; Length 1413;

Best Local Similarity 99.6%; Pred No. 1e-210;

Matches 568; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 345 TGACAAACCAATTCCTCGATGATGCTCTGAGTCTCTGCTGAGGAAACAAATGGGAA 404

DB 1 TGACAAACCAATTCCTCGATGATGCTCTGAGTCTCTGCTGAGGAAACAAATGGGAG 60

QY 405 GTCTGCCAGCAGAGAAAATCTTGCCAAAGCCTTCTGCTCTGCTCAAGGCC 464

DB 61 GCTGCCCCAGCAGAGAAAATCTTGCCAAAGCCTTCTGCTCTGCTCAAGGCC 120

QY 465 AGGACCTGTGACCAATGAG 524

Db 121 AGGACCTGTGACCAATGAG 180

QY 525 TCCCGGCGAGTGGCCCGGCGGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCT 584

Db 181 TCCCGGCGAGTGGCCCGGCGGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCT 240

QY 585 CTGAGGATGAGACTGCTGAGAGCGGTGCTGTGAAGTCTCAGGAGAGAGATTAACATCC 644

Db 241 CTGAGGATGAGACTGCTGAGAGCGGTGCTGTGAAGTCTCAGGAGAGAGATTAACATCC 300

QY 645 CCAGCGTCCACGTGGGCAAGTCTCCATGGGTGCTGTATGAGGGCTGAGCAGGAGAGA 704

Db 301 CCAGCGTCCACGTGGGCAAGTCTCCATGGGTGCTGTATGAGGGCTGAGCAGGAGAGA 360

QY 705 AACGAGAGAACTGCTGTTTACCTGGGAACCTTGGAGGGGCTTCTCTATCCGGGAGA 764

Db 361 AACGAGAGAACTGCTGTTTACCTGGGAACCTTGGAGGGGCTTCTCTATCCGGGAGA 420

QY 765 GCCAGACAGGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCGGCGCTCATCTCTGGG 824

Db 421 GCCAGACAGGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCGGCGCTCATCTCTGGG 480

QY 825 ACCGATCCAGACACTACAGATCCACTGCTTGAACAATGGCTGTACATCTCACCGC 884

Db 481 ACCGATCCAGACACTACAGATCCACTGCTTGAACAATGGCTGTACATCTCACCGC 540

QY 885 GCTCTACCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTG 914

Db 541 GCTCTACCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTG 570

RESULT 9

AAS74748

ID AAS74748 standard; cDNA; 603 BP.

XX AC AAS74748;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #10552.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX FR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG10561.

XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

XX Claim 1; SEQ ID NO 10552; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX SQ Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;  
 Query Match 34.2%; Score 405; DB 5; Length 603;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-181;  
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 779 GGCTCTTACTCTCTGTCAGTCGGCTCAGCCCGCTGATCTCGGACCGGATCAGACAC 838  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 199 GGCTCTTACTCTCTGTCAGTCGGCTCAGCCCGCTGATCTCGGACCGGATCAGACAC 258  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 839 TACAGGATCCATCGCTTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCC 898  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 259 TACAGGATCCATCGCTTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCC 318  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 899 TCACCTCAGCCCTGTGTGACCAATCTACTGTAGCTGGCGGTGACATCTGCTGCTACTC 958  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 319 TCACCTCAGCCCTGTGTGACCAATCTACTGTAGCTGGCGGTGACATCTGCTGCTACTC 378  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 959 AAGAGCCCTGTGTCTGACAGGCGTGGCCCGCTCCCTGGCAAGGATATACCCCTACCT 1018  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 379 AAGAGCCCTGTGTCTGACAGGCGTGGCCCGCTCCCTGGCAAGGATATACCCCTACCT 438  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1019 GTGACTGTGACAGGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTTTCT 1078  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 439 GTGACTGTGACAGGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTTTCT 498  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1079 GAAGCTGCCACAGGAGGAGTCTTCTTCTAGTGAGGTCTCCGGAGTCCCTCAGCTTC 1138  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 499 GAAGCTGCCACAGGAGGAGTCTTCTTCTAGTGAGGTCTCCGGAGTCCCTCAGCTTC 558  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1139 TACATCAGCTGAATCAGCAGGCTGTCTCTTTTGGATGATGCTAG 1183  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 559 TACATCAGCTGAATCAGCAGGCTGTCTCTTTTGGATGATGCTAG 603  
 RESULT 10  
 AAS74750  
 ID AAS74750 standard; cDNA; 2049 BP.  
 XX  
 XX AC AAS74750;  
 XX  
 XX DT 13-FEB-2002 (first entry)  
 XX  
 XX DE DNA encoding novel human diagnostic protein #10554.  
 XX  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX PN W0200175067-A2.  
 XX  
 XX PD 11-OCT-2001.  
 XX  
 XX PF 30-MAR-2001; 2001WO-008631.

XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 XX P-PSDB; ABG10563.  
 PI WPI; 2001-639362/73.  
 XX  
 DR New isolated polynucleotide and encoded polypeptides, useful in  
 DR diagnostics, forensics, gene mapping, identification of mutations  
 DR responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PT  
 XX Claim 1; SEQ ID NO 10554; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX SQ Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;  
 Query Match 33.9%; Score 401; DB 5; Length 2049;  
 Best Local Similarity 99.8%; Pred. No. 4.6e-179;  
 Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 355 ATTTCCCTCCGATGATGCTTCTGAGTGTCTGCTGAGGAACTGGAGTCTGCCAG 414  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 922 ATTTCCCTCCGATGATGCTTCTGAGTGTCTGCTGAGGAACTGGAGTCTGCCAG 981  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 415 CAGAAGAAAATCTTGCAGAGCCCAAGCTTGAATTCCTCTGCTCCAGGCGGACCTGT 474  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 982 CAGAAGAAAATCTTGCAGAGCCCAAGCTTGAATTCCTCTGCTCCAGGCGGACCTGT 1041  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 475 GACCATGGAAGCAGAGAGCAAGCAAGCCAGCCGCTGGCCCTGGGAGTTTCCCGCAGG 534  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1042 GACCATGGAAGCAGAGAGCAAGCAAGCCAGCCGCTGGGAGTTTCCCGCAGG 1101  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 535 TGGCCCGGCGAGCTGTCCCTGAGACTCGGGAGCCATTGACCATCTCTCTGAGATCG 594  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1102 TGGCCCGGCGAGCTGTCCCTGAGACTCGGGAGCCATTGACCATCTCTCTGAGATCG 1161  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 595 AGACTGTGACCGGTGCTGTCTGAAGTCTCAGGACAGAGTATAACATCCCGAGCTCCA 654  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1162 AGACTGTGACCGGTGCTGTCTGAAGTCTCAGGACAGAGTATAACATCCCGAGCTCCA 1221  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 655 CGTGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGACGAGGAGAAAGCAGAGA 714  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1222 CGTGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGACGAGGAGAAAGCAGAGA 1281  
 Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 715 ACTGCTGTGTATTACCTGGGAAACCTCGAGGGGCTTCTCTCATCCGGAGAGCCAGCAG 774  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1282 ACTGCTGTGTATTACCTGGGAAACCTCGAGGGGCTTCTCTCATCCGGAGAGCCAGCAG 1341

QY 775 GAGAGGCTTACTCTGTCTGTCAGTCCGCTCA 806  
 Db 1342 GAGAGGCTTACTCTGTCTGTCAGTCCGCTCA 1373

RESULT 11  
 ABQ98669  
 ID ABQ98669 standard; DNA; 444 BP.

AC ABQ98669;  
 DT 04-NOV-2002 (first entry)  
 DE Human ORF476 coding sequence.

DE Human ORF476 coding sequence.  
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.

OS Homo sapiens.

PN US2002082206-A1.

PD 27-JUN-2002.

PF 30-MAY-2001; 2001US-00867550.

PR 30-MAY-2000; 2000US-0208427P.

PI (LEAC/) LEACH M D.

PA (MEHR/) MEHRABAN F.

PA (CONL/) CONLEY P B.

PA (TOPP/) TOPPER J N.

PA (LAWD/) LAW D.

PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

DR WPI: 2002-626554/67.

DR P-FSDS; ABP64106.

XX New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.

XX Claim 2; SEQ ID NO 951; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?docID=2002082206

XX Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;

Query Match 29.4%; Score 348; DB 6; Length 444;

Best Local Similarity 100.0%; Fred. NO. 5.4e-154;

Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 CCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTCCAG 301

Db 1 CCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTCCAG 60

QY 302 CCAGAGCATGCTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAATTTCC 361  
 Db 61 CCAGAGCATGCTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAATTTCC 120  
 QY 362 TCGATGATGCTCTCAGTGTCTCTCTGAGGACACATGGGAAGTCTGCCGACGACAGA 421  
 Db 121 TCGATGATGCTCTCAGTGTCTCTCTGAGGACACATGGGAAGTCTGCCGACGACAGA 180  
 QY 422 AAATCTCTGCCAAAGCCCAAGCTTGTCTCTCTGTCCTCAAGCCAGGACCTGTGACCATG 481  
 Db 181 AAATCTCTGCCAAAGCCCAAGCTTGTCTCTCTGTCCTCAAGCCAGGACCTGTGACCATG 240  
 QY 482 GAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 541  
 Db 241 GAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
 QY 542 GCCGAGCTGTCTGAGACTCGGGGAGGACATGACCATGCTCTCTGAG 589  
 Db 301 GCCGAGCTGTCTGAGACTCGGGGAGGACATGACCATGCTCTCTGAG 348

RESULT 12

ABQ99151

ID ABQ99151 standard; DNA; 875 BP.

AC ABQ99151;

DT 04-NOV-2002 (first entry)

DE Human ORF958 coding sequence.

DE Human ORF958 coding sequence.  
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.

OS Homo sapiens.

PN US2002082206-A1.

PD 27-JUN-2002.

PF 30-MAY-2001; 2001US-00867550.

PR 30-MAY-2000; 2000US-0208427P.

PI (LEAC/) LEACH M D.

PA (MEHR/) MEHRABAN F.

PA (CONL/) CONLEY P B.

PA (TOPP/) TOPPER J N.

PA (LAWD/) LAW D.

PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

DR WPI: 2002-626554/67.

DR P-FSDS; ABP64588.

XX New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.

XX Claim 2; SEQ ID NO 1915; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological



OS Homo sapiens.  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 XX 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 XX WPI; 2001-639362/73.  
 DR P-PSDB; ABG05994.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 XX Claim 1; SEQ ID NO 5985; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 211 BP; 50 A; 51 C; 72 G; 38 T; 0 U; 0 Other;  
 Query Match 11.9%; Score 141; DB 5; Length 211;  
 Best Local Similarity 99.5%; Pred No. 3.5e-56;  
 Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 588 AGGATGGAGACTGTGGACGGTGTCTCAAGTCTCAGGACAGAGTATACATCCCCA 647  
 DB 20 AGGATGGAGACTGTGGACGGTGTCTCAAGTCTCAGGACAGAGTATACATCCCCA 79  
 QY 648 CGCTCCACGTGGGCAAGTCTCCATCGGGTGTGTATGAGGGCTTACAGCGGAGAAAG 707  
 DB 80 CGCTCCACGTGGGCAAGTCTCCATCGGGTGTGTATGAGGGCTTACAGCGGAGAAAG 139  
 QY 708 CAGAGGAACCTCTTCTTACCTGGGAAACCTCGAGGGGCTTCTTCATCCGGGAGGCC 767  
 DB 140 CAGAGGAACCTCTTCTTACCTGGGAAACCTCGAGGGGCTTCTTCATCCGGGAGGCC 199  
 QY 768 AGACCAGGAGAG 779  
 DB 200 AGACCAGGAGAG 211  
 RESULT 15  
 AA112879  
 ID AA112879 standard; DNA; 432 BP.

XX  
 AC AA112879;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Probe #2812 for gene expression analysis in human cervical cell sample.  
 XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 XX cervical cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000670.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000US-0004263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-488901/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human cervical epithelial cells.  
 PT  
 XX Claim 25; SEQ ID NO 2812; 487pp; English.  
 PS  
 XX The present invention relates to human single exon nucleic acid probes  
 CC (SENPs). The present sequence is one such probe. The SENPs are derived  
 CC from human HeLa cells. The SENPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging of  
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 432 BP; 109 A; 115 C; 111 G; 97 T; 0 U; 0 Other;  
 Query Match 11.3%; Score 134; DB 4; Length 432;  
 Best Local Similarity 100.0%; Pred No. 6.8e-53;  
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 929 GAGCTGGCGGATGACATCTGCTGCTCAAGGAGCCCTGCTCCTGACAGAGGGCTGGC 988  
 DB 270 GAGCTGGCGGATGACATCTGCTGCTCAAGGAGCCCTGCTCCTGACAGAGGGCTGGC 329  
 QY 989 CGCTCCCTCGGCAAGATATACCTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048  
 DB 330 CGCTCCCTCGGCAAGATATACCTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 389  
 QY 1049 AAAGAGCTGGACAG 1062  
 DB 390 AAAGAGCTGGACAG 403  
 Search completed: February 20, 2004, 06:07:58  
 Job time : 512 secs



GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 05:08:33 ; Search time 2929 Seconds

(without alignments)  
12061.099 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccaagacc.....tctcttgatgatgcctag 1183

Scoring table: OLIGO NUC

Gapop\_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estnu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_hic.\*
- 9: gb\_estcl.\*
- 10: gb\_estc2.\*
- 11: gb\_hic.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: em\_ges\_hum.\*
- 18: em\_ges\_inv.\*
- 19: em\_ges\_pln.\*
- 20: em\_ges\_vrt.\*
- 21: em\_ges\_fun.\*
- 22: em\_ges\_mam.\*
- 23: em\_ges\_mus.\*
- 24: em\_ges\_pro.\*
- 25: em\_ges\_rod.\*
- 26: em\_ges\_phg.\*
- 27: em\_ges\_vrl.\*
- 28: gb\_gesl.\*
- 29: gb\_ges2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	735	62.1	1201	9	AL541041
2	647	54.7	1002	12	BQ052308
3	578	48.9	986	12	BQ054265
4	568	48.0	1020	12	BQ054281

RESULT 1  
AL541041  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CG05005YK23  
5-PRIME, mRNA sequence.  
AL541041  
AL541041.2 GI:30544829  
EST.  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (Bases 1 to 1201)  
Li, W.B.; Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:12871733.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 9825.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CG05005AFI2QP1&cluster=9825.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :

# ALIGNMENTS

554	46.8	1069	12	BQ052468
459	38.8	616	13	BX383606
426	36.0	778	12	BG178487
405	34.2	565	12	BG284179
390	33.0	878	12	BQ053486
300	25.4	597	9	AL844311
276	23.3	614	9	AL844307
255	21.5	794	12	BG677567
205	17.3	960	13	BQ944126
200	16.3	606	9	AL844309
116	9.8	842	9	AL844308
106	9.0	611	9	AL844310
103	8.7	506	28	B45150
80	6.8	639	9	AL844312
68	5.7	546	28	AQ556467
57	4.8	149	9	AA309769
47	4.0	553	28	AQ556478
43	3.5	660	10	BG35615
43	3.5	660	13	BY742155
43	3.6	926	11	AK020837
43	3.6	2637	11	AK088672
43	3.6	2974	11	AK030877
42	3.6	322	10	AW437301
42	3.6	569	14	CB426333
34	2.9	377	9	AA959151
30	2.5	389	10	BE015229
29	2.5	603	13	BQ553005
28	2.3	627	10	BG619854
27	2.3	757	29	CC527971
23	1.9	541	12	B1898999
22	1.9	784	12	BG388843
22	1.9	791	14	CB994934
22	1.9	875	12	BG282272
22	1.9	934	12	B1549288
22	1.9	1109	12	BM926459
22	1.9	2517	29	AY418068
22	1.9	2517	29	AY418069
21	1.8	320	14	CB557851
21	1.8	332	10	BE030537
21	1.8	363	9	AA365070
21	1.8	408	29	CG869927

BQ052468 AGENCOURT  
BX383606 BX383606  
BG178487 602328305  
BG284179 602408226  
BQ053486 AGENCOURT  
AL844311 AL844311  
AL844307 AL844307  
BG677567 603624118  
BQ944126 AGENCOURT  
AL844309 AL844309  
AL844308 AL844308  
AL844310 AL844310  
B45150 HS-1060-B1-  
AL844312 AL844312  
AQ556467 HS 5236 B  
AA309769 EST180699  
AQ556478 HS 5236 B  
BG35615 BB35615  
BY742155 BY742155  
AK020837 Mus muscu  
AK088672 Mus muscu  
AK030877 Mus muscu  
AW437301 78366 MAR  
CB426333 601508 MA  
AA959151 vz51g06.r  
BE015229 127457 MA  
BQ553005 H4019E02-  
BB619854 BB619854  
CC527971 CH240\_403  
B1898999 480839 MA  
BG388843 602414563  
CB994934 AGENCOURT  
BG282272 602402951  
B1549288 603189926  
BM926459 AGENCOURT  
AY418068 Homo sapi  
AY418069 Pan trogl  
CB557851 OSJNEC13C  
BE030537 128644 MA  
AA365070 EST175755  
CG869927 XS0853 SA

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DE005AF12QPI.

## FEATURES

Location/Qualifiers  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="CS0DE005YK23"  
/tissue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with NotI and cloned into  
the NotI and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

## ORIGIN

Query Match 62.1%; Score 735; DB 9; Length 1201;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 835; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TAGAGCTCCAGGACCCACGCTGTGTCTCTGTGACAGAGCTCAAGGGCCCTGGGCT 63  
Db 65 TAGAGCTCCAGGACCCACGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGGCT 124

Qy 64 TCCCTCCCTGGCTGGCTGTGCTGGAGGGTTCCCGATCCAGATCCCTAAGAGCAT 123  
Db 125 TCCCTCCCTGGCTGGCTGTGCTGGAGGGTTCCCGATCCAGATCCCTAAGAGCAT 184

Qy 124 GGGGAGCTGATCCATCCCTGGGTGTAACAATCTGACTGCAGACAGATGCTGACTACC 183  
Db 185 GGGGAGCTGATCCATCCCTGGGTGTAACAATCTGACTGCAGACAGATGCTGACTACC 244

Qy 184 CAACCAACACCTAGCTCTCCCTGAGATTCCTCCAGGCTGAGAGAGTTCTGGGTGCC 243  
Db 245 CAACCAACACCTAGCTCTCCCTGAGATTCCTCCAGGCTGAGAGAGTTCTGGGTGCC 304

Qy 244 TAGACCAAGACACTGGGAGACTTCAGAGGGGCCCCCAAGGCTCACTGTCAGGCC 303  
Db 305 TAGACCAAGACACTGGGAGACTTCAGAGGGGCCCCCAAGGCTCACTGTCAGGCC 364

Qy 304 AGACATGCTGCTCAGAGAGCTGTCTTCCAGAGCTTTTGTGATGACAAACCAATTCCTC 363  
Db 365 AGACATGCTGCTCAGAGAGCTGTCTTCCAGAGCTTTTGTGATGACAAACCAATTCCTC 424

Qy 364 GATGATGCTGCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCCCCAGAGAGAA 423  
Db 425 GATGATGCTGCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCCCCAGAGAGAA 484

Qy 424 ATCTCTCCAGCCCAAGCTTGAGTTCTCTGTCAGGCGGAGGACCTGTGACCATGGA 483  
Db 485 ATCTCTCCAGCCCAAGCTTGAGTTCTCTGTCAGGCGGAGGACCTGTGACCATGGA 544

Qy 484 AGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTGGGAGTTTCCCGGAGGTGGCCCGGC 543  
Db 545 AGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTGGGAGTTTCCCGGAGGTGGCCCGGC 604

Qy 544 CGAGCTGCTGCTGAGCTCGGAGAGCATGACCATGCTCTGTGAGATGAGAGTGTG 603  
Db 605 CGAGCTGCTGCTGAGCTCGGAGAGCATGACCATGCTCTGTGAGATGAGAGTGTG 664

Qy 604 GACGGTCTGCTGAGTCTCAGGACAGAGTATACATCCCGAGGCTCCAGTGGGCAA 663  
Db 665 GACGGTCTGCTGAGTCTCAGGACAGAGTATACATCCCGAGGCTCCAGTGGGCAA 724

Qy 664 AGTCTCCAGGGTGGCTGTATCAGGGCTGAGCGGAGAGAGAGAGGAGTCTGTT 723  
Db 725 AGTCTCCAGGGTGGCTGTATCAGGGCTGAGCGGAGAGAGAGAGGAGTCTGTT 784

Qy 724 GTTACCTGGGAACCTTGGAGGGCCCTTCTCATCCCGGAGAGCCAGACAGAGAGGCTC 783  
Db 785 GTTACCTGGGAACCTTGGAGGGCCCTTCTCATCCCGGAGAGCCAGACAGAGAGGCTC 844

Qy

Db

## RESULT 2

BQ052308  
LOCUS  
DEFINITION  
5', mRNA sequence.

VERSION  
BQ052308.1 GI:19811648

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

## REFERENCE

1 (bases 1 to 1002)

## AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgapps-r@mail.nih.gov](mailto:cgapps-r@mail.nih.gov)  
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM2118 row: d column: 23  
High quality sequence stop: 670.

## FEATURES

## source

1..1002  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5933542"  
/tissue\_type="natural killer cells, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH-MGC\_106"  
/note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH-MGC Library."

## ORIGIN

Query Match 54.7%; Score 647; DB 12; Length 1002;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 747; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 333 CCAAGCCTTGATGACAAACCAATTCCTCGATGATGCTCTGAGTCTGCTGCTGAG 392

Db 31 CCAGGCTTTGATGACAAACCAATTCCTCGATGATGCTCTGAGTCTGCTGAG 90

Qy 393 GAACAATGGGAAGTGTGCCAGAGAGAAAATCTTGCCAAAGCCCAAGCTTGAGTTCT 452

Db 91 GAACAATGGGAAGTGTGCCAGAGAGAAAATCTTGCCAAAGCCCAAGCTTGAGTTCT 150

Qy 453 CTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGTGG 512

Db 151 CTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGTGG 210

Qy 513 CCTGGGAGTTTCCCGGAGGTGGCCCGGAGAGTGTGCTGAGACTCGGGAGGCAT 572

Db 211 CCTGGGAGTTTCCCGGAGGTGGCCCGGAGAGTGTGCTGAGACTCGGGAGGCAT 270

Qy 573 TGACCATCTCTCTCAGGATGGAGACTGGTGGACGGTGTCTGTGAAGTCTCAGGCAGAG 632

Db 271 TGACCATCGTCTCTGAGGATGGAGACTGGTGGAGCGGTGCTGTCTGAAGTCTCAGGCAGAG 330  
QY 633 AGTATAACATCCCGAGCGTCCACGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCC 692  
Db 331 AGTATAACATCCCGAGCGTCCACGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCC 390  
QY 693 TGACGAGGAGAAAGCAGAGAACTGCTGTGTGTACTCTGGAAACCTTGAGGGGCTTCC 752  
Db 391 TGAGCAGGAGAAAGCAGAGAACTGCTGTGTGTACTCTGGAAACCTTGAGGGGCTTCC 450  
QY 753 TCATCCGGGAGAGCCAGACAGGAGAGGCTTCTACTCTCTGTGAGTCCCGCTCAGCGGCC 812  
Db 451 TCATCCGGGAGAGCCAGACAGGAGAGGCTTCTACTCTCTGTGAGTCCCGCTCAGCGGCC 510  
QY 813 CTGATCTGGGCGGATGAGCACTACAGGATCCACTCCCTTGACATGCTGCTGT 872  
Db 511 CTGATCTGGGCGGATGAGCACTACAGGATCCACTCCCTTGACATGCTGCTGT 570  
QY 873 ACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCGCTGTGGACCATTAATCTGAGC 932  
Db 571 ACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCGCTGTGGACCATTAATCTGAGC 630  
QY 933 TGGGGATGATCATCTGCTCACTCAAGAGCGCTGTCTGTCAGAGGGTGGCGGCC 992  
Db 631 TGGGGATGATCATCTGCTCACTCAAGAGCGCTGTCTGTCAGAGGGTGGCGGCC 690  
QY 993 TCCCTGGCAAGGATATACCTCTACCTGTGACTGTGAGAGGACCACTCAACTGGAAG 1052  
Db 691 TCCCTGGCAAGGATATACCTCTACCTGTGACTGTGAGAGGACCACTCAACTGGAAG 750  
QY 1053 AGCTGGACAGTCCCTCCTCTTCTGAA 1081  
Db 751 AGCTGGACAGTCCCTCCTCTTCTGAA 779

RESULT 3  
BQ054265  
LOCUS  
DEFINITION AGNCOURT 6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
5', mRNA sequence.

ACCESSION BQ054265  
VERSION BQ054265.1 GI:19813605  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-xemail.nih.gov](mailto:cgabbs-xemail.nih.gov)  
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM2125 row: 1 column: 12  
High quality sequence stop: 515.

FEATURES  
source  
1..986  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5936339"  
/tissue="types=natural killer cells, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_106"  
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

ORIGIN  
Query Match 48.9%; Score 578; DB 12; Length 986;  
Best Local Similarity 99.8%; Pred. No. 1.2e-281;  
Matches 628; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 120 GCATGGGGAGCTGATCCCATCTCCCTGTGTACAACTGCTGACGTCAGACAGATGCTGAGC 179  
Db 1 GCATGGGGAGCTGATCCCATCTCCCTGTGTACAACTGCTGACGTCAGACAGATGCTGAGC 60  
QY 180 TACCCAAACCAACACCTAGCCTCTCCCTGAAGATCCCTCCAGGCTGAGAGAGTCTTGGGT 239  
Db 61 TACCCAAACCAACACCTAGCCTCTCCCTGAAGATCCCTCCAGGCTGAGAGAGTCTTGGGT 120  
QY 240 GTCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTC 299  
Db 121 GTCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTC 180  
QY 300 AGCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTC 359  
Db 181 AGCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTC 240  
QY 360 CCTCGATGATGCTCTGAGTGTCTGCTGAGGACACATGGGAAGTCTGCCAGCAGAA 419  
Db 241 CCTCGATGATGCTCTGAGTGTCTGCTGAGGACACATGGGAAGTCTGCCAGCAGAA 300  
QY 420 GAAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTGTCCAGGCCAGGAGACTGTGACCA 479  
Db 301 GAAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTGTCCAGGCCAGGAGACTGTGACCA 360  
QY 480 TGAAGCAGAGAGAAACCAAGGCCACAGCGCTGCCCTGGGCAGTTCCTCCGGCAGTGGCC 539  
Db 361 TGAAGCAGAGAGAAACCAAGGCCACAGCGCTGCCCTGGGCAGTTCCTCCGGCAGTGGCC 420  
QY 540 CGGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGAGACT 599  
Db 421 CGGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGAGACT 480  
QY 600 GGTGGACGGTGTCTGAGTCTCAGGCAGAGAGTATAACATCCCAAGCTTCCACCTGG 659  
Db 481 GGTGGACGGTGTCTGAGTCTCAGGCAGAGAGTATAACATCCCAAGCTTCCACCTGG 540  
QY 660 GCAAGTCTCCATGGTGGCTGTATGAGGGCTTGAACCATCGTCTCTGAGGATGAGACT 719  
Db 541 CCAAGTCTCCATGGTGGCTGTATGAGGGCTTGAACCATCGTCTCTGAGGATGAGACT 600  
QY 720 TGTGTGTACTCTGGGAACCTCTGGAGGGGCC 748  
Db 601 TGTGTGTACTCTGGGAACCTCTGGAGGGGCC 629

RESULT 4  
BQ054281  
LOCUS  
DEFINITION AGNCOURT 6830234 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936362  
5', mRNA sequence.

ACCESSION BQ054281  
VERSION BQ054281.1 GI:19813621  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2125 row: j column: 11  
High quality sequence stop: 556.  
Location/Qualifiers  
1..1020  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5933772"  
/tissue\_type="natural killer cells, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC\_106"  
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 48.0%; Score 568; DB 12; Length 1020;  
Best Local Similarity 99.6%; Pred. No. 1.5e-276;  
Matches 718; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 191 ACACCTAGCCTTCCTCAAGATCTCCAGGCTGAGAGTCTGGTCTCTAGGACC 250  
DB 5 ACACCTAGCCTTCCTCAAGATCTCCAGGCTGAGAGTCTGGTCTCTAGGACC 64  
QY 251 AAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTCAGCCAGAGCAT 310  
DB 65 AAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTCAGCCAGAGCAT 124  
QY 311 GCGTCTCAGCAGAGCTGCTCCAGCCTTTGATGACAAACCAATTCCTCGATGATG 370  
DB 125 GCGTCTCAGCAGAGCTGCTCCAGCCTTTGATGACAAACCAATTCCTCGATGATG 184  
QY 371 TCGTCTGAGTCTCTGCTGAGGAACAATGGAAAGTCTGCCAGCAGAGAAATCTCTG 430  
DB 185 TCGTCTGAGTCTCTGCTGAGGAACAATGGAAAGTCTGCCAGCAGAGAAATCTCTG 244  
QY 431 CCAAGCCCAAGTTGAGTTCCTCTGTCAGCCAGGACCTGTGACCATGGAGAGAG 490  
DB 245 CCAAGCCCAAGTTGAGTTCCTCTGTCAGCCAGGACCTGTGACCATGGAGAGAG 304  
QY 491 AGAAGCAGAGCCACAGCGTGGCCCTGGGCGAGTTTCCCGCAGGTGGCCCGCAGCTG 550  
DB 305 AGAAGCAGAGCCACAGCGTGGCCCTGGGCGAGTTTCCCGCAGGTGGCCCGCAGCTG 364  
QY 551 TCGCTGAGACTCGGGAGCCATTGACCATGCTCTGAGAGATGAGAGACTGTGGACGGTG 610  
DB 365 TCGCTGAGACTCGGGAGCCATTGACCATGCTCTGAGAGATGAGAGACTGTGGACGGTG 424  
QY 611 CTGCTGAGTCTCAGGACAGAGTATACATCCCGCAGCTCCACGTGGGCAAGTCTCC 670  
DB 425 CTGCTGAGTCTCAGGACAGAGTATACATCCCGCAGCTCCACGTGGGCAAGTCTCC 484  
QY 671 CATGGTGGCTGTATGAGGCTTGAGCAGGGGAGAAAGCAGAGGAACTGCTGTGTTAGCT 730  
DB 485 CATGGTGGCTGTATGAGGCTTGAGCAGGGGAGAAAGCAGAGGAACTGCTGTGTTAGCT 544  
QY 731 GGGAACTTGGAGGGCTTCTCTATCCGGAGAGCCAGACAGAGAGGCTCTTACTCT 790  
DB 545 GNGAACCTTGGAGGGGCTTCTCTATCCGGAGAGCCAGACAGAGAGGCTCTTACTCT 604

QY 791 CTGTCACTCCGCTCAGCCGCTTCATCTCGGACCGGATCAGACTACAGATCCAC 850  
DB 605 CTGTCACTCCGCTCAGCCGCTTCATCTCGGACCGGATCAGACTACAGATCCAC 664  
QY 851 TGCCTTGACATGCTGCTGCTACATCTCAGCGGCTCAGCTTCCCTCACTCCAGGCC 910  
DB 665 TGCCTTGACATGCTGCTGCTACATCTCAGCGGCTCAGCTTCCCTCACTCCAGGCC 724  
QY 911 C 911  
DB 725 C 725

RESULT 5

BQ052468 1069 bp mRNA linear EST 29-MAR-2002  
LOCUS AGENCOURT 6868422 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5933772  
DEFINITION 5', mRNA sequence.

ACCESSION BQ052468  
VERSION BQ052468.1 GI:19811808  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1069)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2118 row: n column: 13  
High quality sequence stop: 681.  
Location/Qualifiers  
1..1069  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5933772"  
/tissue\_type="natural killer cells, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC\_106"  
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

FEATURES

source

ORIGIN

Query Match 46.8%; Score 554; DB 12; Length 1069;  
Best Local Similarity 99.8%; Pred. No. 1.9e-269;  
Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 310 TCGGTCTCAGCAGAGCTGCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCGATGAT 369  
DB 1 TCGGTCTCAGCAGAGCTGCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCGATGAT 60  
QY 370 GTGCTTCTGAGTCTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCAGAGAAATCTCT 429  
DB 61 GTGCTTCTGAGTCTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCAGAGAAATCTCT 120  
QY 430 GCCAAGCCCAAGCTTCTGATTTCTCTCTCCAAAGCCAGGAGCACTGTGACCATGAAGCAGA 489

Db 121 GCCAAGCCCAAGCTTGTCTCTGTCTCAAGCCAGGAGCTGTGACCATGGAAGAGA 180  
 Qy 490 GAGAGCAGGACACAGCGTGGCCCTGGCAGTTTCCCGCAGGTGCGCCGCGGAGCT 549  
 Db 181 GAGAAGCAAGGCGCACAGCGTGGCCCTGGCAGTTTCCCGCAGGTGCGCCGCGGAGCT 240  
 Qy 550 GTCGCTGAGACTCGGGGAGCCATTGACCATCGTCTCTGAGGATGAGACTGTGTGACGGT 609  
 Db 241 GTCGCTGAGACTCGGGGAGCCATTGACCATCGTCTCTGAGGATGAGACTGTGTGACGGT 300  
 Qy 610 GCTGTCTGAGTCTGAGCAGAGATATACATCCCGAGGCTCCACGTGGGCAAGTCTC 669  
 Db 301 GCTGTCTGAGTCTGAGCAGAGATATACATCCCGAGGCTCCACGTGGGCAAGTCTC 360  
 Qy 670 CCATGGTGGCTGTATGAGGGGCTGAGCAGGAGAAAGCAGAGGAATCTGTGTTTACC 729  
 Db 361 CCATGGTGGCTGTATGAGGGGCTGAGCAGGAGAAAGCAGAGGAATCTGTGTTTACC 420  
 Qy 730 TGGGAACCTGAGGGGCTTCTCTATCCGGGAGAGCCAGACGAGGCTCTTACTC 789  
 Db 421 TGGGAACCTGAGGGGCTTCTCTATCCGGGAGAGCCAGACGAGGCTCTTACTC 480  
 Qy 790 TGTGTGAGTCCCGCTCAGCGGCTTCTCTGAGGAGGATCAGACTACAGGATCCA 849  
 Db 481 TGTGTGAGTCCCGCTCAGCGGCTTCTCTGAGGAGGATCAGACTACAGGATCCA 540  
 Qy 850 GTGCTTGAATGAGTGGCTGTATCTACATCCCGGAGGCTTCTCTCTCTCTCTCTCTCT 909  
 Db 541 GTGCTTGAATGAGTGGCTGTATCTACATCCCGGAGGCTTCTCTCTCTCTCTCTCTCT 600  
 Qy 910 CTGGG 914  
 Db 601 CTGGG 605

RESULT 6  
 BX383606  
 LOCUS  
 DEFINITION BX383606 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
 ACCESSION BX383606  
 VERSION BX383606.1 GI:30457152  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 616)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by life technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 9825.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D013BF05QPI&cluster=9825.r. Contact :  
 Peng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D013BF05QPI.  
 Location/Qualifiers  
 1. 516  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D013YK10"  
 /cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
 /cell\_line="JURKAT"  
 /clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT

10-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 ORIGIN  
 Query Match 38.8%; Score 459; DB 13; Length 616;  
 Best Local Similarity 99.8%; Pred. No. 2.7e-221;  
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 9 CTCCAAGGACCCACGCCCTGTCTCTGTGACAGAGCTCAAGGGGCCCTGGGCTTCCCT 68  
 Db 58 CTCCAAGGACCCACGCCCTGTCTCTGTGACAGAGCTCAAGGGGCCCTGGGCTTCCCT 117  
 Qy 69 CCTGTGCTGGGTGTGCTTGGGAGGTTCCCAAGTCCAGAAATCCCTTAAGGAGATGGGGC 128  
 Db 118 CCTGTGCTGGGTGTGCTTGGGAGGTTCCCAAGTCCAGAAATCCCTTAAGGAGATGGGGC 177  
 Qy 129 AGCTGATCCATCCCTGTGTACAAAGTGTGCTGACGACAGATGCTGAGCTACCCAAAC 188  
 Db 178 AGCTGATCCATCCCTGTGTACAAAGTGTGCTGACGACAGATGCTGAGCTACCCAAAC 237  
 Qy 189 CAACACTAGCTCTCTCCCTGAAGATCTCCAGGCTGAGAGATTTGGGTGCTTAGGA 248  
 Db 238 CAACACTAGCTCTCTCCCTGAAGATCTCCAGGCTGAGAGATTTGGGTGCTTAGGA 297  
 Qy 249 CCAAGGACATGACAGACTTCCAGAGGGGCCCCCAAGCCCTAACCTGCCAGCCAGAGC 308  
 Db 298 CCAAGGACATGACAGACTTCCAGAGGGGCCCCCAAGCCCTAACCTGCCAGCCAGAGC 357  
 Qy 309 ATGCGTCTCAGCAGAGCTGTCTTCCCAAGCTTTTGATGACAAACCAATTTCCCTCGATGA 368  
 Db 358 ATGCGTCTCAGCAGAGCTGTCTTCCCAAGCTTTTGATGACAAACCAATTTCCCTCGATGA 417  
 Qy 369 TGTGCTTCTAGTCTCTGCTGAGGACATGGAAGTCTGCCAGCAGAGAAATCTC 428  
 Db 418 TGTGCTTCTAGTCTCTGCTGAGGACATGGAAGTCTGCCAGCAGAGAAATCTC 477  
 Qy 429 TGCCAAAGCCCAAGCTTGAAGTTCCTCTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAG 488  
 Db 478 TGCCAAAGCCCAAGCTTGAAGTTCCTCTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAG 537  
 Qy 489 AGAGAACCAAGGCGACAGCCGTGGCCCTGG 518  
 Db 538 AGAGAACCAAGGCGACAGCCGTGGCCCTGG 567  
 RESULT 7  
 BX178487  
 LOCUS  
 DEFINITION 602328305F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4429896 5',  
 mRNA sequence.  
 ACCESSION BX178487  
 VERSION BX178487.1 GI:12685190  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 778)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: csaps@remail.nih.gov  
 Tissue Procurement: DCTD/DTP  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

BQ053485 878 bp mRNA linear EST 29-MAR-2002  
 AGENCOURT\_6822017 NIH\_MGC\_106 Homo sapiens CDNA clone IMAGE:5935253  
 5', mRNA sequence.  
 BQ053486



```
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 614) Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and
Sheridan, E.
TITLE Homo sapiens EST sequence
JOURNAL Unpublished (2002)
COMMENT Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquerry@sanger.ac.uk
Sanger Centre name : sccdi0816.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool_AK lib v SPD cDNA library. Further information can be found at
http://www.sanger.ac.uk/Teams/Team69/.
FEATURES
Location/Qualifiers
1..614
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="20"
/clone_lib="pool_AK lib v SPD"
ORIGIN
Query Match 23.3%; Score 276; DB 9; Length 614;
Best Local Similarity 100.0%; Pred. No. 2.2e-128;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 908 GCCCTGGTGGACCACTTACTCTGAGTGGCGGATGACATCTGCTGCTACTCAAGAGGCC 967
DB 44 GCCCTGGTGGACCACTTACTCTGAGTGGCGGATGACATCTGCTGCTACTCAAGAGGCC 103
QY 968 TGTGTCTTCGAGAGGCTGGCCCGCTCCCTGGGAGGATATACCCCTACTGTGACTGTG 1027
DB 104 TGTGTCTTCGAGAGGCTGGCCCGCTCCCTGGGAGGATATACCCCTACTGTGACTGTG 163
QY 1028 CAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAAGCTGCC 1087
DB 164 CAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAAGCTGCC 223
QY 1088 ACAGGGAGGAGTCTCTTCTCAGTGGAGTCTCCGGAGTCCCTCAGCTTCTACATCAGC 1147
DB 224 ACAGGGAGGAGTCTCTTCTCAGTGGAGTCTCCGGAGTCCCTCAGCTTCTACATCAGC 283
QY 1148 CTGAATGACGAGGCTGCTCTTTGGATGATGCTAG 1183
DB 284 CTGAATGACGAGGCTGCTCTTTGGATGATGCTAG 319
RESULT 12
BG677567 794 bp mRNA linear EST 01-MAY-2001
DEFINITION 602624118f1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4748884 5',
mRNA sequence.
ACCESSION BG677567
VERSION BG677567.1 GI:13908964
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 794)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
```

```
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAW10601 row: d column: 05
High quality sequence stop: 790.
FEATURES
Location/Qualifiers
1..794
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match 21.6%; Score 255; DB 12; Length 794;
Best Local Similarity 100.0%; Pred. No. 1.1e-117;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 929 GAGCTGGCGGATGACATCTGCTGCTACTCAGGAGCCCTGTCTGCAGAGGCTGCG 988
DB 17 GAGCTGGCGGATGACATCTGCTGCTACTCAGGAGCCCTGTCTGCAGAGGCTGCG 76
QY 989 CCGCTCCCTCCGCAAGATATACCCCTACTGTGACTGTGCAGAGGACACCACTCAACTGG 1048
DB 77 CCGCTCCCTCCGCAAGATATACCCCTACTGTGACTGTGCAGAGGACACCACTCAACTGG 136
QY 1049 AAAGAGCTGGAGAGCTCCCTCTGTTTCTGAAGTCCACAGGGAGAGTCTTCTC 1108
DB 137 AAAGAGCTGGAGAGCTCCCTCTGTTTCTGAAGTCCACAGGGAGAGTCTTCTC 196
QY 1109 AGTGAAGGCTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGCTGTCT 1168
DB 197 AGTGAAGGCTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGCTGTCT 256
QY 1169 TTGGATGATGCTAG 1183
DB 257 TTGGATGATGCTAG 271
RESULT 13
BG944126 960 bp mRNA linear EST 18-OCT-2002
DEFINITION AGENCOURT_10545003 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6728350 5', mRNA sequence.
ACCESSION BG944126
VERSION BG944126.1 GI:24132945
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 960)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ruben Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
```



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM3049 row: m column: 21

High quality sequence stop: 628.

Location/Qualifiers

FEATURES

source

1..960

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6728350"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="PH108 (phage-resistant)"

/clone\_lib="NIH MGC 107"

/note="Organ: Breast; Vector: pOTB7; Site: 1: EcoRI;

Site: 2: XhoI; cDNA made by oligo-dr priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

ORIGIN

Query Match

Best Local Similarity 17.3%; Score 205; DB 13; Length 960;

Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

979 GAGGGTGGCCCGCTCCCTGGCAAGGATATACCCCTGATCTGTCGACAGACACC 1038

DB

177 GAGGGTGGCCCGCTCCCTGGCAAGGATATACCCCTGATCTGTCGACAGACACC 236

QY

1039 ACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGGA 1098

DB

237 ACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGGA 296

QY

1099 GTCTCTTCTCAGTGAGGCTCTCCGGAGTCCTCAGCTTCTACATCAGCTGAATGACGA 1158

DB

297 GTCTCTTCTCAGTGAGGCTCTCCGGAGTCCTCAGCTTCTACATCAGCTGAATGACGA 356

QY

1159 GCGTGTCTCTTTGGATGCGCTAG 1183

DB

357 GCGTGTCTCTTTGGATGCGCTAG 381

RESULT 14

AL844309

LOCUS

AL844309 pool\_FLU\_lib\_v\_SPC Homo sapiens cDNA, mRNA linear EST 30-JUL-2002

DEFINITION

AL844309 pool\_FLU\_lib\_v\_SPC Homo sapiens cDNA, mRNA sequence.

ACCESSION

AL844309.1 GI:22019091

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 608)

Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,B.J. and

Sheridan,E.

Homo sapiens EST sequence

Unpublished (2002)

Contact: The Sanger Centre

The Sanger Centre

Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

Email: humquery@sanger.ac.uk

Sanger Centre name : sccdi0817.154136A

Homo sapiens EST sequence. This sequence was generated as part of

The Wellcome Trust Sanger Institute program to identify and

annotate genes in the human genome. Incomplete or unconfirmed genes

are experimentally analysed using a variety of cDNA library

resources. This sequence was obtained from a PCR product generated

from a pool of up to 100,000 cDNA clones derived from

pool\_FLU\_lib\_v\_SPC cDNA library. Further information can be found

at http://www.sanger.ac.uk/Teams/Team69/.

FEATURES

source

1..606

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/map="20"

/clone\_lib="pool\_FLU\_lib\_v\_SPC"

ORIGIN

Query Match

Best Local Similarity 16.9%; Score 200; DB 9; Length 606;

Matches 250; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

908 GCCTGGTGGACCACTACTGAGCTGGCGGATGACATCTGCTGCTACTCAAGGAGCCC 967

DB

46 GCCTGGTGGACCACTACTGAGCTGGCGGATGACATCTGCTGCTACTCAAGGAGCCC 105

QY

968 TGTGTCTCTGAGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACTGTGACTGTG 1027

DB

106 TGTGTCTCTGAGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACTGTGACTGTG 165

QY

1028 CAGAGGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCC 1087

DB

166 CAGAGGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCC 225

QY

1088 ACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGC 1147

DB

226 ACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGC 285

QY

1148 CTGAATGACGA 1158

DB

286 CTGAATGACGA 296

RESULT 15

AL844308/c

LOCUS

AL844308 pool\_AK\_lib\_v\_SPD Homo sapiens cDNA, mRNA linear EST 30-JUL-2002

DEFINITION

AL844308 pool\_AK\_lib\_v\_SPD Homo sapiens cDNA, mRNA sequence.

ACCESSION

AL844308.1 GI:22019090

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 642)

Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,B.J. and

Sheridan,E.

Homo sapiens EST sequence

Unpublished (2002)

Contact: The Sanger Centre

The Sanger Centre

Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

Email: humquery@sanger.ac.uk

Sanger Centre name : sccdi0816.400489A

Homo sapiens EST sequence. This sequence was generated as part of

The Wellcome Trust Sanger Institute program to identify and

annotate genes in the human genome. Incomplete or unconfirmed genes

are experimentally analysed using a variety of cDNA library

resources. This sequence was obtained from a PCR product generated

from a pool of up to 100,000 cDNA clones derived from

pool\_AK\_lib\_v\_SPD cDNA library. Further information can be found at

http://www.sanger.ac.uk/Teams/Team69/.

Location/Qualifiers

1..642

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/map="20"

/clone\_lib="pool\_AK\_lib\_v\_SPD"

FEATURES

source

1..642

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/map="20"

/clone\_lib="pool\_AK\_lib\_v\_SPD"

ORIGIN

Query Match 9.8%; Score 116; DB 9; Length 642;  
 Best Local Similarity 100.0%; Pred.No. 3.7e-47;  
 Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1068 TCCTGTTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGT 1127  
 DB 588 TCCTGTTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGT 529  
 QY 1128 CCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCCTAG 1183  
 DB 528 CCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCCTAG 473

Search completed: February 20, 2004, 08:13:51  
 Job time : 2937 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 05:30:56 ; Search time 113 Seconds  
(without alignments)  
5809.796 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagactccaagacc.....tctcttgatgagcctag 1183

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/2/ina/5A-COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B-COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/6A-COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/6B-COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PCITUS-COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	1.7	539	US-09-621-976-10381	Sequence 10381, A
2	20	1.7	675	US-08-707-793A-3	Sequence 3, Appli
3	20	1.7	675	US-08-707-792A-3	Sequence 3, Appli
4	20	1.7	2129	US-09-016-434-1452	Sequence 1452, Ap
5	20	1.7	786431	US-09-751-388-3	Sequence 3, Appli
6	19	1.6	1467	US-09-579-182-2	Sequence 2, Appli
7	19	1.6	1548	US-09-099-053-1	Sequence 1, Appli
8	19	1.6	2771	US-09-016-434-1101	Sequence 1101, Ap
9	18	1.5	1438	US-09-187-331-4	Sequence 4, Appli
10	18	1.5	1438	US-09-470-946-4	Sequence 4, Appli
11	18	1.5	1569	US-09-318-448-8	Sequence 8, Appli
12	18	1.5	3090	US-09-276-531-78	Sequence 78, Appli
13	18	1.5	70000	US-09-851-898-3	Sequence 3, Appli
14	17	1.4	351	US-09-046-479-1	Sequence 1, Appli
15	17	1.4	351	US-08-822-897C-1	Sequence 1, Appli
16	17	1.4	351	US-09-608-810A-3	Sequence 3, Appli
17	17	1.4	351	US-09-404-417A-1	Sequence 1, Appli
18	17	1.4	435	US-09-252-991A-6817	Sequence 6817, Ap
19	17	1.4	439	US-09-222-575-172	Sequence 172, App
20	17	1.4	439	US-09-389-681-172	Sequence 172, App
21	17	1.4	439	US-09-620-405B-172	Sequence 172, App
22	17	1.4	439	US-09-339-338-172	Sequence 172, App
23	17	1.4	439	US-09-433-826B-172	Sequence 172, App
24	17	1.4	439	US-09-604-287A-172	Sequence 172, App
25	17	1.4	439	US-09-285-480-172	Sequence 172, App
26	17	1.4	439	US-09-834-759-172	Sequence 172, App
27	17	1.4	445	US-09-702-705-1598	Sequence 1598, Ap

28	17	1.4	445	US-09-736-457-1598	Sequence 1598, Ap
29	17	1.4	445	US-09-614-124B-1598	Sequence 1598, Ap
30	17	1.4	445	US-09-671-325-1598	Sequence 1598, Ap
31	17	1.4	541	US-09-220-132-10	Sequence 10, Appl
32	17	1.4	566	US-09-621-976-1574	Sequence 1574, Ap
33	17	1.4	627	US-09-328-352-1086	Sequence 1086, Ap
34	17	1.4	643	US-09-833-381-1262	Sequence 1262, Ap
35	17	1.4	651	US-09-016-434-1255	Sequence 1255, Ap
36	17	1.4	674	US-09-621-976-87	Sequence 87, Appl
37	17	1.4	759	US-09-489-039A-5493	Sequence 5493, Ap
38	17	1.4	894	US-09-434-354-1	Sequence 1, Appli
39	17	1.4	941	US-09-634-238-89	Sequence 89, Appl
40	17	1.4	1089	US-08-154-915-1	Sequence 1, Appli
41	17	1.4	1089	US-08-464-517-37	Sequence 37, Appl
42	17	1.4	1089	US-08-246-361A-37	Sequence 37, Appl
43	17	1.4	1089	US-08-463-772-37	Sequence 37, Appl
44	17	1.4	1089	PCT-US93-09945-1	Sequence 1, Appli
45	17	1.4	1747	US-09-566-921-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-10381  
; Sequence 10381, Application US/09621976  
; Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S. Y.

FILE OF INVENTION: ESTS and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 10381

LENGTH: 539

TYPE: DNA

ORGANISM: Homo sapiens

US-09-621-976-10381

Query Match 1.7%; Score 20; DB 4; Length 539;

Best Local Similarity 100.0%; Pred. No. 4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 472 TGTGACCATGGAAGCAGAGA 491

Db 52 TGTGACCATGGAAGCAGAGA 71

RESULT 2

US-08-707-793A-3  
; Sequence 3, Application US/08707793A  
; Patent No. 5776696

GENERAL INFORMATION:

APPLICANT: SALOWE, SCOTT P.

TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING

TITLE OF INVENTION: FUSION PROTEINS

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESS: Merck & Co., Inc.

STREET: P.O. Box 2000, 126 E. Lincoln Ave.

CITY: Rahway

STATE: NJ

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/707,793A  
;; FILING DATE: 04-SEP-1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Camara, Valerie J  
;; REGISTRATION NUMBER: 35,090  
;; REFERENCE/DOCKET NUMBER: 19494  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 908-594-3902  
;; TELEFAX: 908-594-4720  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 675 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: Genomic DNA  
US-08-707-793A-3

Query Match 1.7%; Score 20; DB 1; Length 675;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766  
Db 443 CCTTCTCATCCGGGAGGC 462

## RESULT 3

US-08-707-792A-3  
;; Sequence 3, Application US/08707792A  
;; Patent No. 5783398  
;; GENERAL INFORMATION:  
;; APPLICANT: MARCY, ALICE  
;; APPLICANT: SALOME, SCOTT P.  
;; APPLICANT: WISNIEWSKI, DOUGLAS  
;; TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING  
;; FUSION PROTEINS  
;; NUMBER OF SEQUENCES: 17  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Merck & Co., Inc.  
;; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
;; CITY: Rahway  
;; STATE: NJ  
;; COUNTRY: USA  
;; ZIP: 07065-0900  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/707,792A  
;; FILING DATE: 04-SEP-1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Camara, Valerie J  
;; REGISTRATION NUMBER: 35,090  
;; REFERENCE/DOCKET NUMBER: 19524  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 908-594-3902  
;; TELEFAX: 908-594-4720  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 675 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: Genomic DNA  
US-08-707-792A-3

Query Match 1.7%; Score 20; DB 1; Length 675;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766  
Db 443 CCTTCTCATCCGGGAGGC 462

## RESULT 4

US-09-016-434-1452  
;; Sequence 1452, Application US/09016434  
;; Patent No. 6500938  
;; GENERAL INFORMATION:  
;; APPLICANT: Janice Au-Young  
;; APPLICANT: Jeffrey J. Seilhamer  
;; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
;; PATHWAY GENE EXPRESSION  
;; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
;; NUMBER OF SEQUENCES: 1490  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
;; STREET: 3174 PORTER DRIVE  
;; CITY: PALO ALTO  
;; STATE: CALIFORNIA  
;; COUNTRY: USA  
;; ZIP: 94304  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/016,434  
;; FILING DATE: HEREWITH  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Zeller, Karen J.  
;; REGISTRATION NUMBER: 37,071  
;; REFERENCE/DOCKET NUMBER: PA-0002 US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (650) 855-0555  
;; TELEFAX: (650) 845-4166  
;; INFORMATION FOR SEQ ID NO: 1452:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2129 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; LIBRARY: GENBANK  
;; CLONE: 9775207  
US-09-016-434-1452

Query Match 1.7%; Score 20; DB 4; Length 2129;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766  
Db 508 CCTTCTCATCCGGGAGGC 527

```
RESULT 5
US-09-751-389-3
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01067
; CURRENT APPLICATION NUMBER: US/09/751.389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match 1.7%; Score 20; DB 4; Length 786431;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 TGAGCAGGAGGAGGAGCAG 712
DB 412751 TGAGCAGGAGGAGGAGCAG 412770

RESULT 6
US-09-579-182-2
; Sequence 2, Application US/09579182
; Patent No. 6500628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MNI-161
; CURRENT APPLICATION NUMBER: US/09/579.182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-579-182-2

Query Match 1.6%; Score 19; DB 4; Length 1467;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 AGGGGCTTCCTCATCCGG 760
DB 423 AGGGGCTTCCTCATCCGG 441

RESULT 7
US-09-099-053-1
; Sequence 1, Application US/09099053
; Patent No. 6388063
; GENERAL INFORMATION:
; APPLICANT: Greg Plowman
; APPLICANT: Susan Onrust
; APPLICANT: David Markby
; APPLICANT: Sara Courtneidge
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: SAD RELATED DISORDERS
; NUMBER OF SEQUENCES: 28
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099.053
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/049,914
; FILING DATE: June 18, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 235/121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1548 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-099-053-1

Query Match 1.6%; Score 19; DB 4; Length 1548;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 AGGGGCTTCCTCATCCGG 760
DB 471 AGGGGCTTCCTCATCCGG 489

RESULT 8
US-09-016-434-1101
; Sequence 1101, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION/DOCKET NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2771 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: G1256002  
US-09-016-434-1101

Query Match 1.6%; Score 19; DB 4; Length 2771;  
Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;  
Matches 19; Conservative 0;

QY 595 CTGAGGATGGAGACTGGTG 603  
DB 1305 CTGAGGATGGAGACTGGTG 1323

RESULT 9  
US-09-187-331-4  
; Sequence 4, Application US/09187331  
; Patent No. 6043056  
; GENERAL INFORMATION:  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Gorgone, Gina A.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS  
; FILE REFERENCE: PF-0631 US  
; CURRENT APPLICATION NUMBER: US/09/187,331  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 1438  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 2705267  
US-09-187-331-4

Query Match 1.5%; Score 18; DB 3; Length 1438;  
Best Local Similarity 100.0%; Pred. No. 38; Mismatches 0; Indels 0; Gaps 0;  
Matches 18; Conservative 0;

QY 1139 TACATCAGCCTGAATGAC 1156  
DB 855 TACATCAGCCTGAATGAC 872

RESULT 10  
US-09-470-946-4  
; Sequence 4, Application US/09470946  
; Patent No. 6358923  
; GENERAL INFORMATION:  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Gorgone, Gina A.  
; APPLICANT: Baughn, Mariah R.

; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS  
; FILE REFERENCE: PF-0631 US  
; CURRENT APPLICATION NUMBER: US/09/470,946  
; CURRENT FILING DATE: 1999-12-22  
; EARLIER APPLICATION NUMBER: US 09/187,331  
; EARLIER FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 1438  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 2705267  
US-09-470-946-4

Query Match 1.5%; Score 18; DB 4; Length 1438;  
Best Local Similarity 100.0%; Pred. No. 38; Mismatches 0; Indels 0; Gaps 0;  
Matches 18; Conservative 0;

QY 1139 TACATCAGCCTGAATGAC 1156  
DB 855 TACATCAGCCTGAATGAC 872

RESULT 11  
US-09-318-448-8  
; Sequence 8, Application US/09318448  
; Patent No. 6210950  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, William G.  
; APPLICANT: Stenroos, Edward S.  
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS  
; FILE REFERENCE: 601-1-057  
; CURRENT APPLICATION NUMBER: US/09/318,448  
; CURRENT FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 1669  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-318-448-8

Query Match 1.5%; Score 18; DB 3; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 38; Mismatches 0; Indels 0; Gaps 0;  
Matches 18; Conservative 0;

QY 416 AGAAGAAAATCTCTGCCA 433  
DB 494 AGAAGAAAATCTCTGCCA 511

RESULT 12  
US-09-276-531-78/c  
; Sequence 78, Application US/09276531  
; Patent No. 6183968  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Lal, Preeti  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Yue, Henry  
; APPLICANT: Reddy, Roopa  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING  
; TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION  
; NUMBER OF SEQUENCES: 134  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO

STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/276,531  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/079,677  
FILING DATE: March 27, 1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: LYNN E. MURRY, Ph.D.  
REGISTRATION NUMBER: 42,918  
REFERENCE/DOCKET NUMBER: PA-0008 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3090 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAINOT14  
CLONE: 1595762  
US-09-276-531-78

Query Match 1.5%; Score 18; DB 3; Length 3090;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1093 GGAGGAGTCTTCTTCAG 1110  
Db 398 GGAGGAGTCTTCTTCAG 381

RESULT 13  
US-09-851-896-3  
Sequence 3, Application US/09851896  
Patent No. 6410325  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Susan M. Freier  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)  
FILE REFERENCE: RTS-0220  
CURRENT APPLICATION NUMBER: US/09/851,896  
CURRENT FILING DATE: 2001-05-08  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 3  
LENGTH: 70000  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-851-896-3

Query Match 1.5%; Score 18; DB 4; Length 70000;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 455 GTCCAGGCCGAGGACCT 472  
Db 60708 GTCCAGGCCGAGGACCT 60725

RESULT 14  
US-09-046-479-1/c  
Sequence 1, Application: US/09046479  
Patent No. 6291653  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Deisher, Theresa A.  
TITLE OF INVENTION: MOTILIN HOMOLOGS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESS: ZymoGenetics, Inc.  
STREET: 1201 Eastlake Avenue East  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,479  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sawislak, Deborah A.  
REGISTRATION NUMBER: 37,438  
REFERENCE/DOCKET NUMBER: 97-04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-442-6672  
TELEFAX: 206-442-6678  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...351  
OTHER INFORMATION:  
NAME/KEY: sig\_peptide  
LOCATION: 1...69  
OTHER INFORMATION:  
NAME/KEY: mat\_peptide  
LOCATION: 70...351  
OTHER INFORMATION:  
US-09-046-479-1

Query Match 1.4%; Score 17; DB 3; Length 351;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 296 GTCCAGGCCGAGCATGC 312  
Db 57 GTCCAGGCCGAGCATGC 41

RESULT 15  
US-08-822-897C-1/c  
Sequence 1, Application US/08822897C  
Patent No. 6380158  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Deisher, Theresa A.  
TITLE OF INVENTION: MOTILIN HOMOLOGS  
NUMBER OF SEQUENCES: 7

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,897C
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sawislak, Deborah A
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 97-04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...351
OTHER INFORMATION:
NAME/KEY: sig_peptide
LOCATION: 1...69
OTHER INFORMATION:
NAME/KEY: mat_peptide
LOCATION: 70...351
OTHER INFORMATION:
US-08-822-897C-1

```

```

Query Match 1.4%; Score 17; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 296 GTCCAGCCAGCATGC 312
Db 57 GTCCAGCCAGCATGC 41

```

```

Search completed: February 20, 2004, 08:16:02
Job time : 121 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 07:26:02 ; Search time 486 Seconds  
(without alignments)

8523.550 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccagagacc.....tctcttgatgacgctag 1183

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2308684 seqs, 1750822206 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	735	62.1	786	14	US-10-043-649-1
2	724	61.2	864	10	US-09-814-353-21302
3	657	55.5	763	9	US-09-867-550-953
4	348	29.4	444	9	US-09-867-550-951
5	341	28.8	875	9	US-09-867-550-1915
6	134	11.3	432	9	US-09-864-761-2829
7	134	11.3	448	9	US-09-864-761-15513
8	96	8.1	96	9	US-09-864-761-19612
9	87	7.4	320	10	US-09-814-353-17314
10	77	6.5	152	10	US-09-814-353-4631
11	71	6.0	152	10	US-09-814-353-10930
12	21	1.8	701	15	US-10-062-674-2188
13	20	1.7	611	15	US-10-027-632-195852
14	20	1.7	672	15	US-10-027-632-107077
15	20	1.7	672	15	US-10-027-632-142058

16	20	1.7	672	15	US-10-027-632-142059	Sequence 142059,
17	20	1.7	672	15	US-10-027-632-142060	Sequence 142060,
18	20	1.7	711	15	US-10-027-632-26286	Sequence 26286, A
19	20	1.7	934	15	US-10-260-238-640	Sequence 640, App
20	20	1.7	2017	15	US-10-062-674-1776	Sequence 1776, Ap
21	20	1.7	2032	15	US-10-366-288-27	Sequence 27, Appl
22	20	1.7	2129	10	US-09-360-706-954	Sequence 954, App
23	20	1.7	2129	15	US-10-305-720-1452	Sequence 1452, Ap
24	20	1.7	786431	14	US-10-412-277-3	Sequence 3, Appl1
25	19	1.6	65	10	US-09-908-975-4510	Sequence 4510, Ap
26	19	1.6	114	9	US-09-864-761-30106	Sequence 30106, A
27	19	1.6	510	14	US-10-072-602B-237	Sequence 237, App
28	19	1.6	599	9	US-09-864-761-13565	Sequence 13565, A
29	19	1.6	1033	9	US-09-789-561-20	Sequence 20, Appl
30	19	1.6	1125	15	US-10-027-632-118578	Sequence 118578,
31	19	1.6	1133	12	US-10-424-599-55347	Sequence 55347, A
32	19	1.6	1467	14	US-10-354-358-11	Sequence 11, Appl
33	19	1.6	1580	10	US-09-976-782-25	Sequence 25, Appl
34	19	1.6	1833	9	US-09-861-846-1	Sequence 1, Appl1
35	19	1.6	2120	15	US-10-094-749-577	Sequence 577, App
36	19	1.6	2771	15	US-10-305-720-1101	Sequence 1101, Ap
37	19	1.6	4207	15	US-10-369-493-46381	Sequence 46381, A
38	19	1.6	4720	15	US-10-062-674-2048	Sequence 2048, Ap
39	18	1.5	204	9	US-09-738-626-3261	Sequence 3261, Ap
40	18	1.5	402	10	US-09-918-995-8609	Sequence 8609, Ap
41	18	1.5	403	9	US-09-983-965-1815	Sequence 1815, Ap
42	18	1.5	474	15	US-10-027-632-195991	Sequence 195991,
43	18	1.5	487	10	US-09-918-995-26739	Sequence 26739, A
44	18	1.5	491	10	US-09-518-995-38019	Sequence 38019, A
45	18	1.5	497	15	US-10-027-632-270409	Sequence 270409,

#### ALIGNMENTS

RESULT 1  
US-10-043-649-1  
; Sequence 1, Application US/10043649  
; Publication No. US2003005924A1  
; GENERAL INFORMATION:  
; APPLICANT: Holland, Sacha J.  
; APPLICANT: Mendenhall, Marcy K.  
; APPLICANT: Pardo, Jorge  
; APPLICANT: Spencer, Collin  
; APPLICANT: Fu, C. Alan  
; APPLICANT: Luo, Ying  
; APPLICANT: Payan, Donald G.  
; APPLICANT: Mancebo, Helena S.Y.  
; APPLICANT: Wu, Jun  
; APPLICANT: Zhou, Xiulan  
; APPLICANT: Shen, Mary  
; APPLICANT: Liao, X. Charlene  
; APPLICANT: Sheng, Ning  
; TITLE OF INVENTION: Cloning of a No. US2003005924A1el Inhibitor of Antigen-receptor  
; FILE REFERENCE: A-70219-1/RMS/DHR  
; CURRENT APPLICATION NUMBER: US/10/043,649  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/260,953  
; PRIOR FILING DATE: 2001-01-10  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 786  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(786)  
; OTHER INFORMATION:  
US-10-043-649-1

Query Match 62.1%; Score 735; DB 14; Length 786;

```

Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGAGTCTGCCAGAGAGAAATCTCTGCAAGCCCAAGCTGAGTCTCTGTC 457
Db 1 ATGGAGTCTGCCAGAGAGAAATCTCTGCAAGCCCAAGCTGAGTCTCTGTC 60
QY 458 CAAGCCAGGAGCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 517
Db 61 CAAGCCAGGAGCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 120
QY 518 GGCAGTTTCCCGGAGGTGGCCCGGAGTGTCTGAGACTCGGGAGCCATTGACC 577
Db 121 GGCAGTTTCCCGGAGGTGGCCCGGAGTGTCTGAGACTCGGGAGCCATTGACC 180
QY 578 ATCGTCTCTGAGATGGAGACTGTGTGACGCTGTCTGAGTCTCAGGCAGAGATAT 637
Db 181 ATCGTCTCTGAGATGGAGACTGTGTGACGCTGTCTGAGTCTCAGGCAGAGATAT 240
QY 638 AACATCCCGAGCTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 697
Db 241 AACATCCCGAGCTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 300
QY 698 AGGAGAAAGCAGAGAACTCTGTGTGACCTGGGACCTGGGGGCTTCTCTATC 757
Db 301 AGGAGAAAGCAGAGAACTCTGTGTGACCTGGGAACTGGAGGGGCTTCTCTATC 360
QY 758 CGGGAGAGCCAGCAGAGAGGCTTCTACTCTGTGTCAGTCCGCTCAGCCGCTGCA 817
Db 361 CGGGAGAGCCAGCAGAGAGGCTTCTACTCTGTGTCAGTCCGCTCAGCCGCTGCA 420
QY 818 TCCTGGGACCGATCAGACACTACAGATCCACTGCTTGAATGCTGCTGATC 877
Db 421 TCCTGGGACCGATCAGACACTACAGATCCACTGCTTGAATGCTGCTGATC 480
QY 878 TCACGCGGCTCACCTTCCCTCTACTCCAGGCGCTGTGGAGCACTTACTCTGAGCTGGC 937
Db 481 TCACGCGGCTCACCTTCCCTCTACTCCAGGCGCTGTGGAGCACTTACTCTGAGCTGGC 540
QY 938 GATGATCTGTGCTACTCAGAGGCGCTGTCTGTCAGAGGGCTGGCCGCTGCT 997
Db 541 GATGATCTGTGCTACTCAGAGGCGCTGTCTGTCAGAGGGCTGGCCGCTGCT 600
QY 998 GCAAGGATATACCCCTACCTGTGACTGTGAGAGGACCACTCAACTGGAAGAGCTG 1057
Db 601 GCAAGGATATACCCCTACCTGTGACTGTGAGAGGACCACTCAACTGGAAGAGCTG 660
QY 1058 GACAGTCCCTCTGTTTGAAGCTGCCAGGGAGGAGTCTCTTCTCAGTGAGGGT 1117
Db 661 GACAGTCCCTCTGTTTGAAGCTGCCAGGGAGGAGTCTCTTCTCAGTGAGGGT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGAGGCTGTCTCTTTGATGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGAGGCTGTCTCTTTGATGAT 780
QY 1178 GCCTAG 1183
Db 781 GCCTAG 786

```

RESULT 2

```

US-09-814-353-21302
; Sequence 21302, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353

```

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21302
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 32, 862, 863, 864
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21302

Query Match 61.2%; Score 724; DB 10; Length 864;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTAGAGCTCCAGAGACCCACGCTGTCTCTGTGACAGAGCTCAAGAGGCCCTGGGC 61
Db 54 GCTAGAGCTCCAGAGACCCACGCTGTCTCTGTGACAGAGCTCAAGAGGCCCTGGGC 113
QY 62 CTTCCCTCCCTGGCTGGTGGAGGGTTCCTCCAGTCCAGATCCCTTAAGAGAGC 121
Db 114 CTTCCCTCCCTGGCTGGTGGAGGGTTCCTCCAGTCCAGATCCCTTAAGAGAGC 173
QY 122 ATGGGACGCTGATCATCCCTGGTGTACAACTGCTGCTGACAGAGATCTGAGCTA 181
Db 174 ATGGGACGCTGATCATCCCTGGTGTACAACTGCTGCTGACAGAGATCTGAGCTA 233
QY 182 CCCAAACCAACACCTAGCTCTCCCTGAGAGATCTCCAGAGGCTGAGAGAGTCTGGGT 241
Db 234 CCCAAACCAACACCTAGCTCTCCCTGAGAGATCTCCAGAGGCTGAGAGAGTCTGGGT 293
QY 242 CCTAGACCAAGAGACTGGAGACTTCCAGAGGCCCCCAAGCCCTTAAGCTTCCAG 301
Db 294 CCTAGACCAAGAGACTGGAGACTTCCAGAGGCCCCCAAGCCCTTAAGCTTCCAG 353
QY 302 CCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGTGACAAACCAATTTCC 361
Db 354 CCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGTGACAAACCAATTTCC 413
QY 362 TCGATGATGCTTCTGAGTGTCTGAGAGAACTGGGAGTCTGCCAGAGAGAGA 421
Db 414 TCGATGATGCTTCTGAGTGTCTGAGAGAACTGGGAGTCTGCCAGAGAGAGA 473
QY 422 AAATCTCTGCCAAGCCCAAGCTTGTGATCTCTGTCCCAAGCCCTTGTGACCAATG 481
Db 474 AAATCTCTGCCAAGCCCAAGCTTGTGATCTCTGTCCCAAGCCCTTGTGACCAATG 533
QY 482 GAAGCAGAGAGAGAGAGCCACAGCCGTGGCCCTTCCCGCAGTTCCTCCGAGTGGCCG 541
Db 534 GAAGCAGAGAGAGAGAGCCACAGCCGTGGCCCTTCCCGCAGTTCCTCCGAGTGGCCG 593
QY 542 GCGAGCTGTCTGAGACTCGGGAGCCATGACATCGTCTCTGAGGATGAGAGCTGG 601
Db 594 GCGAGCTGTCTGAGACTCGGGAGCCATGACATCGTCTCTGAGGATGAGAGCTGG 653
QY 602 TGAACGGTGTCTGAGAGTCTCAGCAGAGAGTATACATCCCGCAGCTCCACGTGGGC 661
Db 654 TGAACGGTGTCTGAGAGTCTCAGCAGAGAGTATACATCCCGCAGCTCCACGTGGGC 713

```

QY 662 AAAGTCTCCATGGGTGGCTGTATGAGGGCTTGAGGAGGAGAAAGCAGAGGAAGTGTCTG 721  
Db 714 AAAGTCTCCATGGGTGGCTGTATGAGGGCTTGAGGAGGAGAAAGCAGAGGAAGTGTCTG 773  
QY 722 TTGTTACTCGGGAACCCCTGGAGGGGCTTCTCATCCGGAGAGCCAGACCCAGGA 776  
Db 774 TTGTTACTCGGGAACCCCTGGAGGGGCTTCTCATCCGGAGAGCCAGACCCAGGA 828

## RESULT 3

US-09-867-550-953  
; Sequence 953, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 953  
; LENGTH: 763  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-867-550-953

Query Match 55.5%; Score 657; DB 9; Length 763;  
Best Local Similarity 99.7%; Pred. No. 0;

Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 117 GGAGCATGGGAGCTGATCCATCCCTGGTGACAACTGCTGACTGAGACAGATGCTG 176  
Db 5 GGAGCATGGGAGCTGATCCATCCCTGGTGACAACTGCTGACTGAGACAGATGCTG 64  
QY 177 AGCTACCCAAACCAACACCTGACCTCTCCCTGAAGATCTCCAGGCTGAGAGGTTCTG 236  
Db 65 AGCTACCCAAACCAACACCTGACCTCTCCCTGAAGATCTCCAGGCTGAGAGGTTCTG 124  
QY 237 GGTGTCCTAGGACCAAGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACTG 296  
Db 125 GATGTCCTAGGACCAAGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACTG 184  
QY 297 TCCAGCCAGAGCATGGTCTCAGCAGAGCTGTCTTCCCAAGCTTTGATGACAAACCAAT 356  
Db 185 TCCAGCCAGAGCATGGTCTCAGCAGAGCTGTCTTCCCAAGCTTTGATGACAAACCAAT 244  
QY 357 TTCCCTCGATGATGCTTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCA 416  
Db 245 TTCCCTCGATGATGCTTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCA 304  
QY 417 GAAGAAATCTCTGCCAAGCCCAAGCTTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCA 476  
Db 305 GAAGAAATCTCTGCCAAGCCCAAGCTTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCA 364  
QY 477 CCATGGAAGCAGAGAGAGCAAGCCACAGCCCTGGCCCTGGCAGGTTTCCCGGCGAGTG 536  
Db 365 CCATGGAAGCAGAGAGAGCAAGCCACAGCCCTGGCCCTGGCAGGTTTCCCGGCGAGTG 424  
QY 537 GCCCGGCGAGTGTCGTGAGCTCGGGAGCCATTGACCATGCTCTGAGATGGAG 596  
Db 425 GCCCGGCGAGTGTCGTGAGCTCGGGAGCCATTGACCATGCTCTGAGATGGAG 484  
QY 597 ACTGTTGAGACGGTGTCTGTAAGTCTCAGGACAGAGTATAACATCCCAAGCTGCCACG 656  
Db 485 ACTGTTGAGACGGTGTCTGTAAGTCTCAGGACAGAGTATAACATCCCAAGCTGCCACG 544

QY 657 TGGCCAAAGTCTCCATGGGTGGCTGTATGAGGGCTTGAGGAGGAGAAAGCAGAGGAAC 716  
Db 545 TGGCCAAAGTCTCCATGGGTGGCTGTATGAGGGCTTGAGGAGGAGAAAGCAGAGGAAC 604  
QY 717 TGCTGTTGTTACTTGGGAACCCCTGGAGGGGCTTCTCATCCGGAGAGCCAGACCCAGGA 776  
Db 605 TGCTGTTGTTACTTGGGAACCCCTGGAGGGGCTTCTCATCCGGAGAGCCAGACCCAGGA 664  
QY 777 GAGGCTCTTACTCTCTCTCAGTCCGCTCAGCCGCTTGCATCTCTGGGACCGGATCAGAC 836  
Db 665 GAGGCTCTTACTCTCTCTCAGTCCGCTCAGCCGCTTGCATCTCTGGGACCGGATCAGAC 724  
QY 837 ACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACA 875  
Db 725 ACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACA 763

## RESULT 4

US-09-867-550-951  
; Sequence 951, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 951  
; LENGTH: 444  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-867-550-951

Query Match 29.4%; Score 348; DB 9; Length 444;  
Best Local Similarity 100.0%; Pred. No. 7.5e-167;

Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 CCTAGCACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGGCCCTTAACCTGTCCAG 301  
Db 1 CCTAGCACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGGCCCTTAACCTGTCCAG 60  
QY 302 CCAGAGATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGTATGACAAACCAATTTCCC 361  
Db 61 CCAGAGATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGTATGACAAACCAATTTCCC 120  
QY 362 TCATGATGTCTTCTGAGTGTCTCTCAGAGAAACAATGGGAAGTCTGCCAGCAGAGA 421  
Db 121 TCATGATGTCTTCTGAGTGTCTCTCTGAGAAACAATGGGAAGTCTGCCAGCAGAGA 180  
QY 422 AATCTCTGCCAAGCCCAAGCTTGTATTCCTGTCTTCCAAAGCCAGGACCTGTGACCATG 481  
Db 181 AATCTCTGCCAAGCCCAAGCTTGTATTCCTGTCTTCCAAAGCCAGGACCTGTGACCATG 240  
QY 482 GAAGCAGAGAGAGCAAGCCACAGCCGTGGCCCTGGGAGTTTCCCGCAGGTGGCCCG 541  
Db 241 GAAGCAGAGAGAGCAAGCCACAGCCGTGGCCCTGGGAGTTTCCCGCAGGTGGCCCG 300  
QY 542 GCGAGCTCTCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAG 589  
Db 301 GCGAGCTCTCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAG 348

## RESULT 5

```

US-09-867-550-1915
; Sequence 1915, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; FILE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1915
; LENGTH: 875
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: Wherein n is one of a or t or c or g
US-09-867-550-1915

Query Match      28.8%; Score 341; DB 9; Length 875;
Best Local Similarity 100.0%; Pred. No. 2.5e-163; Indels 0; Gaps 0;
Matches 341; Conservative 0; Mismatches 0;

QY 843 GGATCCACTGCTTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCTCAC 902
Db 2 GGATCCACTGCTTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCTCAC 61
QY 903 TCCAGGCCCTGGTGACCAATTAATCTGAGCTGGCGGAGTACATCTGTGCTACTCAAG 962
Db 62 TCCAGGCCCTGGTGACCAATTAATCTGAGCTGGCGGAGTACATCTGTGCTACTCAAG 121
QY 963 AGCCCTGTCTCTGTCAGAGGCTGCGCGCTCCCTGGCAAGATATACCCCTACTCTGA 1022
Db 122 AGCCCTGTCTCTGTCAGAGGCTGCGCGCTCCCTGGCAAGATATACCCCTACTCTGA 181
QY 1023 CTGTGACAGACACCACTCACTGGAAGAGCTGGAGCTGCTCTCTGTTTCTGAAG 1082
Db 182 CTGTGACAGACACCACTCACTGGAAGAGCTGGAGCTGCTCTCTGTTTCTGAAG 241
QY 1083 CTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGTCTCCGGAGTCCCTCAGCTTCTACA 1142
Db 242 CTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGTCTCCGGAGTCCCTCAGCTTCTACA 301
QY 1143 TCAGCCTGAATGACAGGCTGTCTCTTTGGAGTATGCTAG 1183
Db 302 TCAGCCTGAATGACAGGCTGTCTCTTTGGAGTATGCTAG 342

RESULT 6
US-09-864-761-2829
; Sequence 2829, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

```

```

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Arcomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2829
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031662.24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ST474, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
US-09-864-761-2829

Query Match      11.3%; Score 134; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 8.3e-58;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 929 GAGCTGGCGGATGACATCTGCTGCTCTCAAGAGCCCTGTCTCTGAGAGGCTGGC 988
Db 270 GAGCTGGCGGATGACATCTGCTGCTCTCAAGAGCCCTGTCTCTGAGAGGCTGGC 329
QY 989 CCGCTCCCTGGCAAGATATACCCCTACTGTGATGTGCGAGGACACCCTCAACTGG 1048
Db 330 CCGCTCCCTGGCAAGATATACCCCTACTGTGATGTGCGAGGACACCCTCAACTGG 389
QY 1049 AAGAGCTGGACAG 1062
Db 390 AAGAGCTGGACAG 403

RESULT 7
US-09-864-761-15513

```

DB 406 AAAGAGCTGGACAG 419

|||||

RESULT 8

US-03-864-761-19612

Sequence 19612, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aecmica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 19612

LENGTH: 96

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AL03162.24

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.4

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3

OTHER INFORMATION: NT HIT: AF000716.1, EVALUATE 1.70e-01

```

: ORGANISM: Homo sapiens
:
: FEATURE:
:
: OTHER INFORMATION: MAP TO AL031662.24
:
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
:
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
:
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.4
:
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
:
: OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.1
:
: OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.9
:
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
:
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
:
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
:
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
:
: OTHER INFORMATION: NT HIT: AF000716.1, EVALUAE 1.70e-01
:

```

```
; OTHER INFORMATION: EST_HUMAN HIT: A1125308.1, EVALUE 2.10e-01
US-09-864-761-19612

Query Match
Best Local Similarity 8.1%; Score 96; DB 9; Length 96;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 967 CTGTCTCTCGAGAGGCTGGCCCTCCCTGGCAAGATATACCCCTACCTGTGACTGT 1026
Db 1 CTGTCTCTCGAGAGGCTGGCCCTCCCTGGCAAGATATACCCCTACCTGTGACTGT 60

QY 1027 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 1062
Db 61 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 96

RESULT 9
US-09-814-353-17314
; Sequence 17314, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17314
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-17314

Query Match
Best Local Similarity 7.4%; Score 87; DB 10; Length 320;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGGTGACGG 608
Db 103 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGGTGACGG 162

QY 609 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGGTGACGG 635
Db 163 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGGTGACGG 189

RESULT 10
US-09-814-353-4631
; Sequence 4631, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
```

```
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4631
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 17, 102, 112
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-4631

Query Match
Best Local Similarity 6.5%; Score 77; DB 10; Length 152;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGGTGACGG 608
Db 25 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGGTGACGG 84

QY 609 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGGTGACGG 625
Db 85 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGGTGACGG 101

RESULT 11
US-09-814-353-10930
; Sequence 10930, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10930
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 17, 102, 112
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-10930

Query Match
Best Local Similarity 6.5%; Score 77; DB 10; Length 152;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 TGTGCTGAGACTCGGGAGCCATTGACCATCGTCTGAGGATGGAGACTGTTGACGG 608
Db 25 TGTGCTGAGACTCGGGAGCCATTGACCATCGTCTGAGGATGGAGACTGTTGACGG 84

QY 609 TGTGCTGAGACTCGGGAGCCATTGACCATCGTCTGAGGATGGAGACTGTTGACGG 625
Db 85 TGTGCTGAGACTCGGGAGCCATTGACCATCGTCTGAGGATGGAGACTGTTGACGG 101

RESULT 12
US-10-062-674-2188/c
; Sequence 2188, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 2188
; LENGTH: 701
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040005559A1 893157.1
; NAME/KEY: unsure (701)
; LOCATION: (1) ...
; OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-2188

Query Match
Best Local Similarity 1.8%; Score 21; DB 15; Length 701;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 TGGCCCTGGGCGAGTTTCCCGG 530
Db 285 TGGCCCTGGGCGAGTTTCCCGG 265

RESULT 13
US-10-027-632-195852/c
; Sequence 195852, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107077
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-195852

Query Match
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 TGGCCCTGGGCGAGTTTCCCGG 530
Db 285 TGGCCCTGGGCGAGTTTCCCGG 265

RESULT 13
US-10-027-632-195852/c
; Sequence 195852, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107077
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-195852

Query Match
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 TGGCCCTGGGCGAGTTTCCCGG 530
Db 285 TGGCCCTGGGCGAGTTTCCCGG 265
```

```
;
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195852
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-195852

Query Match
Best Local Similarity 1.7%; Score 20; DB 15; Length 611;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 CACTGGCAGACTTCCAGAAG 275
Db 503 CACTGGCAGACTTCCAGAAG 484

RESULT 14
US-10-027-632-107077
; Sequence 107077, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107077
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-107077

Query Match
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766
Db 71 CCTTCTCATCCGGGAGGC 90

RESULT 15
US-10-027-632-142058
; Sequence 142058, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 142058

; LENGTH: 672

; TYPE: DNA

; ORGANISM: Human

; US-10-027-632-142058

Query Match 1.7%; Score 20; DB 15; Length 672;

Best Local Similarity 100.0%; Pred. No. 9.5;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CTTCTCATCCGGGAGAGC 766

Db 71 CTTCTCATCCGGGAGAGC 90

Search completed: February 20, 2004, 09:27:51

Job time : 487 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 08:14:02 ; Search time 93 Seconds

(without alignments)

792.956 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLPGRRKSLPSPSLSSV.....RESLSFVSLNDEAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1353	100.0	261	5	AAU91308 Human pro
2	1347	99.6	261	5	AAO15457 Human mod
3	1347	99.6	261	5	ABP52187 Human Src
4	1347	99.6	261	5	AAE26357 Human Src
5	1269	93.8	248	3	AAE42993 Human ORF
6	1028	76.0	259	5	AAO15456 Mouse mod
7	934.5	69.1	210	5	AAO15458 Mouse mod
8	923.5	68.3	210	5	ABP64788 Human pro
9	826	61.0	159	5	ABP64107 Human ORF
10	586	43.3	113	5	ABP64588 Human ORF
11	481.5	35.6	276	5	ABP52188 Human Src
12	481.5	35.6	276	6	ABU79103 Src-like
13	481.5	35.6	276	6	ABU59698 Human Src
14	470.5	34.8	315	4	AAU31072 Novel hum
15	452.5	33.4	96	5	ABP64106 Human ORF
16	370.5	27.4	512	7	ADD19014 Human dis
17	360.5	26.6	505	4	AAE99332 Human tyr
18	360.5	26.6	505	7	ABWQ1407 Human hae
19	348	25.7	509	3	AAU49420 PKA subst
20	348	25.7	509	6	ABU58699 Human can
21	348	25.7	509	6	ABU58699 Human can
22	340	25.1	363	6	ABU59690 Human pro
23	340	25.1	437	5	ABG79672 Tumour in
24	340	25.1	508	3	ABU37700 Human lym
25	340	25.1	508	7	ABU58802 Human pro

26	340	25.1	508	7	ABU58799 Human pro
27	340	25.1	509	7	ABU58799 Human pro
28	338	25.0	70	4	ABG05994 Novel hum
29	337	24.9	567	5	ABG79673 Tumour in
30	332	24.5	517	4	ABU57957 Drosophila
31	317	23.4	541	5	AAU74614 Perinucle
32	315.5	23.3	536	7	ADD45918 Human pro
33	315.5	23.3	537	7	ADD14129 Human src
34	315.5	23.3	543	4	AAAB4663 Amino aci
35	315.5	23.3	543	4	ABG10302 Novel hum
36	315.5	23.3	543	6	ADA00843 Human Src
37	313.5	23.2	496	2	AAU29658 Human src
38	313.5	23.2	496	4	AAU08734 Xenopus 1
39	313.5	23.2	496	4	AAU08730 Xenopus 1
40	313.5	23.2	496	4	AAU08735 Xenopus 1
41	313.5	23.2	543	2	AAU24421 Human yes
42	312.5	23.1	529	7	ABG31767 Human 145
43	306.5	22.7	551	4	ABG22264 Novel hum
44	303	22.4	458	7	ABG99048 Human kpp
45	286.5	21.2	533	3	AAU44449 Mutant ch

## ALIGNMENTS

### RESULT 1

AAU91308

ID AAU91308 standard; protein; 261 AA.

XX AC AAU91308;

XX DT 18-JUN-2002 (first entry)

XX DE Human protein NOV13.

XX KW Human; NOVX; gene therapy; cardiomyopathy; atherosclerosis;

XX KW cell signal processing disorder; metabolic pathway modulation disorder;

XX KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;

XX KW uterus cancer; immune response; graft-versus-host disease;

XX KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;

XX KW hypertension; congenital heart defects; multiple sclerosis; inflammation;

XX KW Albright hereditary osteodystrophy.

XX OS Homo sapiens.

XX PN WO200216599-A2.

XX PD 28-FEB-2002.

XX PF 27-AUG-2001; 2001WO-US026510.

XX PR 25-AUG-2000; 2000US-0228191P.

XX PR 08-FEB-2001; 2001US-0267300P.

XX PR 20-FEB-2001; 2001US-0269961P.

XX PR 20-MAR-2001; 2001US-0277337P.

XX PA (CURA-) CURAGEN CORP.

XX PA (CORT-) COR THERAPEUTICS INC.

XX PI Burgess CB, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;

XX PI Spytek KA, Szekeres BS, Tomlinson JE, Topper JN, Yang R;

XX DR WPI: 2002-280937/32.

XX DR N-PSDB; ABK61465.

XX PT New polypeptides for treating or preventing a disorder associated with

XX PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.

XX PS Claim 3; Page 98; 263pp; English.

XX CC The invention relates to an isolated polypeptide (NOVX) a mature form of

XX CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide

XX CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,

15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it and antibody against it, are useful for treating or preventing (e.g. by gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy, atherosclerosis, a disorder related to cell signal processing and metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide and nucleic acids are also useful for determining the presence of a predisposition to the disease. The NOVX nucleic acid and polypeptide are especially useful in therapeutic or prophylactic applications for disorders associated with aberrant NOVX expression or activity, e.g. cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus cancer), immune response, graft-versus-host disease, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension, congenital heart defects, multiple sclerosis, inflammation or Albritght hereditary osteodysrophy and many other diseases listed in the specification. The DNA encoding the protein is useful in gene therapy for treating the conditions. This is also useful in detection assays, chromosome mapping, tissue typing, diagnostic or prognostic assays, or for developing a powerful assay system for functional analysis of various human disorders, as well as in diagnostic applications. The present sequence represents a NOVX protein

XX Sequence 261 AA;

Query Match 100.0%; Score 1353; DB 5; Length 261;  
Best Local Similarity 100.0%; Pred. No. 1.6e-130;  
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSLSRKRKSLPSPSSVSGQGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
DB 1 MGSLSRKRKSLPSPSSVSGQGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60

QY 61 IVSEGDGDMWTVLSEVSGREYNIPSVHVGVKSHGWLVEGLSREKAEELLLPGNPGGAFLI 120  
DB 61 IVSEGDGDMWTVLSEVSGREYNIPSVHVGVKSHGWLVEGLSREKAEELLLPGNPGGAFLI 120

QY 121 RESQTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFPSLOALVDHYSELA 180  
DB 121 RESQTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFPSLOALVDHYSELA 180

QY 181 DDICLLKEPCVLQAGPLPGKDIPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEG 240  
DB 181 DDICLLKEPCVLQAGPLPGKDIPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEG 240

QY 241 LRESLSFYISLNDEAVSLDDA 261  
DB 241 LRESLSFYISLNDEAVSLDDA 261

RESULT 2

AAO15457  
ID AAO15457 standard; protein; 261 AA.

XX AAO15457;

XX 03-OCT-2002 (first entry)

XX Human modulator of antigen receptor signalling (MARS) protein.

XX Human; gene therapy; modulator of antigen receptor signalling; MARS;  
KW tumour suppressor gene; Src-like adaptor protein; SLAP;  
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
KW immunosuppression; myeloproliferative disorder; breast cancer.

OS Homo sapiens.

XX WO200242452-A2.

XX 30-MAY-2002.

XX 26-NOV-2001; 2001WO-CA001662.

XX 27-NOV-2000; 2000CA-02324663.

XX

PA (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX McGlade JC, Loreto MP;

XX WPI; 2002-566564/60.

XX N-PSDB; AAL44089.

XX New isolated modulator of antigen receptor signaling protein or its fragment, useful for treating malignant disorders such as myeloid malignancies, autoimmune disorders and myeloproliferative disorders.

XX Claim 7; Fig 9A; 110pp; English.

XX The invention comprises the amino acid and coding sequences of modulator of antigen receptor signalling (MARS) proteins. The MARS protein is a putative tumour suppressor gene and exhibits structural and sequence similarity to the Src-like adaptor protein (SLAP). The MARS DNA and protein sequences of the invention are useful for the treatment of myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune disorders, immunosuppression, myeloproliferative disorders and malignancies related to the de-regulation of tyrosine kinases (e.g. breast cancer). The present amino acid sequence represents a human MARS protein

XX Sequence 261 AA;

Query Match 99.6%; Score 1347; DB 5; Length 261;  
Best Local Similarity 99.6%; Pred. No. 6.8e-130;  
Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSLSRKRKSLPSPSSVSGQGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
DB 1 MGSLSRKRKSLPSPSSVSGQGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60

QY 61 IVSEGDGDMWTVLSEVSGREYNIPSVHVGVKSHGWLVEGLSREKAEELLLPGNPGGAFLI 120  
DB 61 IVSEGDGDMWTVLSEVSGREYNIPSVHVGVKSHGWLVEGLSREKAEELLLPGNPGGAFLI 120

QY 121 RESQTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFPSLOALVDHYSELA 180  
DB 121 RESQTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFPSLOALVDHYSELA 180

QY 181 DDICLLKEPCVLQAGPLPGKDIPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEG 240  
DB 181 DDICLLKEPCVLQAGPLPGKDIPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEG 240

QY 241 LRESLSFYISLNDEAVSLDDA 261  
DB 241 LRESLSFYISLNDEAVSLDDA 261

RESULT 3

ABP52187

ID ABP52187 standard; protein; 261 AA.

XX ABP52187;

XX 15-OCT-2002 (first entry)

XX Human Src-like inhibitory molecule (SLIM) protein.

XX Human; Src-like inhibitory molecule; SLIM; Src-like adaptor protein;  
KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;  
KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;  
KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
KW chronic inflammatory disorder; autoimmune disorder; transplant rejection.

XX Homo sapiens.

XX WO200255707-A2.

XX 18-JUL-2002.

XX

PF 10-JAN-2002; 2002WO-US000718.  
 XX  
 PR 10-JAN-2001; 2001US-0260953P.  
 XX  
 PA (RIGE-) RIGEL PHARM INC.  
 XX  
 PI Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
 XX  
 DR WPI; 2002-575432/61.  
 DR N-PSDB; ABQ/4343.  
 XX  
 PT New src-like inhibitory molecule protein, useful for treating  
 PT immunodeficiency disorders and inflammatory disorders, comprises N-  
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.  
 XX  
 PS Claim 7; Fig 2A; 91pp; English.  
 XX  
 CC The present sequence represents the human Src-like inhibitory molecule  
 CC (SLIM) protein (I). The present invention describes a SLIM protein  
 CC comprising an N-terminal myristylation sequence, an N-terminal SH2  
 CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising  
 CC an N-terminal myristylation sequence and an N-terminal SH2 domain which  
 CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and  
 CC anti-Hiv activities, and can be used as a modulator of lymphocyte  
 CC activation, and of ubiquitination of a Cbl target protein, and in gene  
 CC therapy. (I) is useful for screening a bioactive agent capable of binding  
 CC to SLIM. (I) is also useful for screening a bioactive agent capable of  
 CC modulating SLIM binding. (I) or its fragments is useful in the study or  
 CC in the treatment of conditions which involves this function or  
 CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or  
 CC prevent SLIM associated disorders. (I) or the polynucleotide encoding it  
 CC (II) is useful for modulating leukocyte and/or platelet activation, for  
 CC modulating antigen receptor-induced signalling and activation in  
 CC leukocyte and/or platelet and for modulating antigen receptor-induced  
 CC signalling and activation in lymphocytes and/or mast cells. (I) or (II)  
 CC is also useful for modulating the basal activity of lymphocytes. (I) or  
 CC (II) is useful in the treatment of immunodeficiency disorders, such as  
 CC acquired immunodeficiency syndrome (AIDS), for the prevention and  
 CC treatment of acute inflammatory disorders, chronic inflammatory  
 CC disorders, autoimmune disorder and transplant rejection  
 XX  
 SQ Sequence 261 AA;  
 Query Match 99.6%; Score 1347; DB 5; Length 261;  
 Best Local Similarity 99.6%; Pred. No. 6.8e-130;  
 Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGSLPFRKSLPSLSVSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 Db 1 MGSLPFRKSLPSLSVSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 QY 61 IVSEGDWMTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKABELLLLPNPGGAFLI 120  
 Db 61 IVSEGDWMTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKABELLLLPNPGGAFLI 120  
 QY 121 RRSQTRRGYSLSVLRSPASWDRIHVRHICLDNGWLYISPLTFPSLQALVCHYSELA 180  
 Db 121 RRSQTRRGYSLSVLRSPASWDRIHVRHICLDNGWLYISPLTFPSLQALVCHYSELA 180  
 QY 181 DPICLLKEPCVLQAGLPFGKDIPLVTVQRTPLNWKELDSLLFSAATGESLLSEG 240  
 Db 181 DPICLLKEPCVLQAGLPFGKDIPLVTVQRTPLNWKELDSLLFSAATGESLLSEG 240  
 QY 241 LRESLSFYLSINDEAVSLDDA 261  
 Db 241 LRESLSFYLSINDEAVSLDDA 261  
 RESULT 4  
 ID AAE26357  
 XX AAE26357 standard; protein; 261 AA.

AC AAE26357;  
 XX 13-DEC-2002 (first entry)  
 DT  
 DE Human Src-Like Adapter Protein-2 (hSLAP-2).  
 XX  
 KW Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 35..90  
 FT /note= "SH3 domain"  
 FT Modified-site 80  
 FT /note= "Tyrosine phosphorylation site"  
 FT Domain 94..176  
 FT /note= "SH2 domain"  
 FT Modified-site 96  
 FT /note= "Tyrosine phosphorylation site"  
 FT Modified-site 130  
 FT /note= "Tyrosine phosphorylation site"  
 FT Modified-site 148  
 FT /note= "Tyrosine phosphorylation site"  
 FT Modified-site 159  
 FT /note= "Tyrosine phosphorylation site"  
 FT Modified-site 176  
 FT /note= "Tyrosine phosphorylation site"  
 PN WO200242457-A1.  
 XX 30-MAY-2002.  
 XX 20-NOV-2001; 2001WO-US043367.  
 XX 22-NOV-2000; 2000US-0252545P.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;  
 XX WPI; 2002-463632/49.  
 DR N-PSDB; AAD43980.  
 XX  
 XX Novel substantially purified human SH2/SH3-domain-containing adapter  
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic  
 PT intervention in immunological and inflammatory disorders and cancer.  
 XX  
 PS Claim 9; Fig 2; 85pp; English.  
 XX  
 CC The invention relates to a substantially purified human SH2/SH3-domain-  
 CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-  
 CC 2). The invention is useful for treating an immune disorder involving  
 CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is  
 CC useful for screening for antagonists or inhibitors of the interaction of  
 CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or  
 CC preventing diseases or disorders associated with aberrant or uncontrolled  
 CC cellular signal transduction, for determining those cellular signalling  
 CC molecules which associate with hSLAP-2 and which provide critical signals  
 CC for cell activation, and as effectors in methods to affect T- cell  
 CC activation. The invention is useful in screening assays to identify and  
 CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for  
 CC potential use to treat autoimmune diseases which may be caused by  
 CC hyperactivated B cells, as well as to treat diseases which may be caused  
 CC by hyperactivated T cells, in addition to other immune system related  
 CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,  
 CC inflammation disorders, diseases and conditions, rheumatoid arthritis,  
 CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's

CC and ulcerative colitis), allergies, particularly those involving  
 CC hyperactivity of B-cells and T-cells, or other immune cells, such as  
 CC mast cells or eosinophils, autoimmune diseases such as systemic lupus  
 CC erythematosus and multiple sclerosis, pulmonary diseases including  
 CC asthma, acute respiratory distress syndrome, and chronic obstructive  
 CC pulmonary disorder, tissue/organ rejection and cancer. The invention is  
 CC useful in gene therapy. The present sequence is human SLAP-2 protein  
 XX  
 XX Sequence 261 AA;  
 Query Match 99.6%; Score 1347; DB 5; Length 261;  
 Best Local Similarity 99.6%; Pred. No. 6.8e-130;  
 Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGSLSRRKSLPSLSLSSVQGGPVTMEABRSKATAVALGSPAGGPAELSLRLGPLET 60  
 Db 1 MGSLSRRKSLPSLSLSSVQGGPVTMEABRSKATAVALGSPAGGPAELSLRLGPLET 60  
 QY 61 IVSEGDWMTVLSEVSGREYNTPSVHVGVKSHGWLVEGLSREKABELLLPFGGAPLI 120  
 Db 61 IVSEGDWMTVLSEVSGREYNTPSVHVGVKSHGWLVEGLSREKABELLLPFGGAPLI 120  
 QY 121 RESQTRRGYSLSVRLSPASWDRIHRYHICLDNGWLYISPLTFPSLQALVDHYSELA 180  
 Db 121 RESQTRRGYSLSVRLSPASWDRIHRYHICLDNGWLYISPLTFPSLQALVDHYSELA 180  
 QY 181 DDIICLLKEPCVLRAGPLPGKDIPLPVTQRTPLNWKELSSLLFSAATGEESLLSEG 240  
 Db 181 DDIICLLKEPCVLRAGPLPGKDIPLPVTQRTPLNWKELSSLLFSAATGEESLLSEG 240  
 QY 241 LRESLSFYISLNDVAVSLDDA 261  
 Db 241 LRESLSFYISLNDVAVSLDDA 261

RESULT 5

AAAB42993  
 ID AAB42993 standard; protein; 248 AA.  
 XX  
 AC AAB42993;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF2757 polypeptide sequence SEQ ID NO:5514.  
 XX  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnery; antiporiatic; antiparkinsonian; nortropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US008621.  
 XX  
 PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX

PA (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Leach M;  
 XX WPI; 2000-602362/57.  
 DR N-PSDB; AAC77202.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 11; Page 4693-4694; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 CC antiporiatic; antiparkinsonian; nortropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 XX Sequence 248 AA;  
 Query Match 93.8%; Score 1269; DB 3; Length 248;  
 Best Local Similarity 93.8%; Pred. No. 6.8e-122;  
 Matches 244; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 15 SLSSSVQGGPVTMEABRSKATAVALGSPAGGPAELSLRLGPLETIVSEGDWMTVLS 74  
 Db 2 SLSSSVQGGPVTMEABRSKATAVALGSPAGGPAELSLRLGPLETIVSEGDWMTVLS 61  
 QY 75 VSGREYNIPSVHVGVKSHGWLVEGLSREKABELLLPFGGAPLIRESQTRGYSLSV 134  
 Db 62 VSGREYNIPSVHVGVKSHGWLVEGLSREKABELLLPFGGAPLIRESQTRGYSLSV 121  
 QY 135 RLSPASWDRIHRYHICLDNGWLYISPLTFPSLQALVDHYSELADDCCLLKEPCVLIQ 194  
 Db 122 RLSPASWDRIHRYHICLDNGWLYISPLTFPSLQALVDHYSELADDCCLLKEPCVLIQ 181  
 QY 195 RAGPLPGKDIPLPVTQRTPLNWKELSSLLFSAATGEESLLSEGLRSLSFYISLND 254  
 Db 182 RAGPLPGKDIPLPVTQRTPLNWKELSSLLFSAATGEESLLSEGLRSLSFYISLND 241  
 QY 255 AVSLDDA 261  
 Db 242 AVSLDDA 248  
 C  
 RESULT 6  
 AAO15456  
 ID AAO15456 standard; protein; 259 AA.  
 XX  
 AC AAO15456;  
 XX  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Mouse modulator of antigen receptor signalling (MARS) protein.  
 XX  
 KW Mouse; gene therapy; modulator of antigen receptor signalling; MARS;  
 KW tumour suppressor gene; Scr-like adaptor protein; SLAP;

KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 OS Mus sp.  
 FN WO200242452-A2.  
 PD 30-MAY-2002.  
 XX 26-NOV-2001; 2001WO-CA001662.  
 PF 27-NOV-2000; 2000CA-02324663.  
 XX (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 PA Mcglade JC, Loreto MP;  
 PI WPI; 2002-566564/60.  
 XX N-PSDB; AAL44087.  
 XX New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX Claim 7; Fig 1A; 110pp; English.  
 XX The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present amino acid sequence represents a mouse MARS  
 CC protein  
 XX Sequence 259 AA;  
 SQ Query Match 76.0%; Score 1028; DB 5; Length 259;  
 Best Local Similarity 79.4%; Pred. No. 4.9e-97;  
 Matches 208; Conservative 16; Mismatches 34; Indels 4; Gaps 3;  
 QY 1 MGSILPSRRKSLPSPSLSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 DB 1 MGSILSSRGKT-SSPSPSSSGPDDQEPVSMQPERHKVATAVALGSPFAGGPAELSLRLGEPLT 59  
 QY 61 IVSEDDGWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLLPGNPGGAFLI 120  
 DB 60 IISEDGDWTVQVEVSGREYHMSVYVAKVAHGLVEGLSREKAEELLLLPGNPGGAFLI 119  
 QY 121 RESQTRRGYSLSVLRSPASWDRIHRYRHCLDNGWLYISPLTTPFSLOALVDHYSELA 180  
 DB 120 RESQTRRGYSLSVLRSPASWDRIHRYRHCLDNGWLYISPLTTPFSLOALVDHYSELA 179  
 QY 181 DDICLLKEPCVLQACPLKQDIPVTVQRTPLNWKELDSSLFSEA-ATGEESLLSE 239  
 DB 180 DGICCPRECEVQLKGLPGLKTPPPVTVFTSSLNWKKLDRLSLFLFEPASGEASLLSE 239  
 QY 240 GLRESLSFYSLNDEAVSLDDA 261  
 DB 240 GLRESLSFYSLNDEAVSLDDA 259  
 RESULT 7  
 ID AAO15458 standard; protein; 210 AA.  
 AC AAO15458;  
 XX 03-OCT-2002 (first entry)  
 DT Mouse modulator of antigen receptor signalling short isoform protein.  
 DE

XX Mouse; gene therapy; modulator of antigen receptor signalling; MARS;  
 KW tumour suppressor gene; Scr-like adaptor protein; SLAP;  
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 XX Mus sp.  
 OS WO200242452-A2.  
 FN 30-MAY-2002.  
 XX 26-NOV-2001; 2001WO-CA001662.  
 PF 27-NOV-2000; 2000CA-02324663.  
 XX (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 PA Mcglade JC, Loreto MP;  
 PI WPI; 2002-566564/60.  
 XX N-PSDB; AAL44090.  
 XX New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX Claim 8; Page 78; 110pp; English.  
 XX The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present amino acid sequence represents a mouse MARS  
 CC protein  
 XX Sequence 210 AA;  
 SQ Query Match 69.1%; Score 934.5; DB 5; Length 210;  
 Best Local Similarity 88.9%; Pred. No. 1.5e-87;  
 Matches 184; Conservative 3; Mismatches 13; Indels 7; Gaps 1;  
 QY 1 MGSILPSRRKSLPSPSLSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 DB 1 MGSILPSRRKSLPSPSLSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 QY 61 IVSEDDGWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLLPGNPGGAFLI 120  
 DB 61 IVSEDDGWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLLPGNPGGAFLI 120  
 QY 121 RESQTRRGYSLSVLRSPASWDRIHRYRHCLDNGWLYISPLTTPFSLOALVDHYSELA 178  
 DB 121 RESQTRRGYSLSVLRSPASWDRIHRYRHCLDNGWLYISPLTTPFSLOALVDHYSELA 180  
 QY 179 ----LADDICLLKEPCVLQACPLP 200  
 DB 181 PAPQGYTPTCDCAEDTTQLERAGQLP 207  
 RESULT 8  
 ID ABP64788 standard; protein; 210 AA.  
 AC ABP64788;  
 XX 25-FEB-2003 (first entry)  
 DT Human protein SEQ ID 448.  
 XX

Human; expressed sequence tag; EST; haematopoietic disorder;  
 central nervous system disease; viral infection;  
 peripheral nervous system disease; non-healing wound; infectious disease;  
 immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
 fungal infection; autoimmune disorder; coagulation disorder; neutropenic;  
 antiallergic; antiinflammatory; immunosuppressive; neuroprotective;  
 cytostatic; haemostatic; virucide; antibacterial; fungicide;  
 immunostimulant; cerebroprotective.

Homo sapiens.

WO200259260-A2.

01-AUG-2002.

16-NOV-2001; 2001WO-US042950.

17-NOV-2000; 2000US-00714936.

(HYSE-) HYSEQ INC.

Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
 Ren F, Xue AU, Yang Y, Wehrman T, Drmanac RT;  
 WPI; 2002-590824/63.  
 N-PSDB; ABO99374.

New isolated polynucleotide, useful in research, diagnostic or  
 therapeutic methods, e.g. preventing or treating disorders involving  
 aberrant protein expression or biological activity.

Claim 20; SEQ ID NO 448; 394pp; English.

The present invention relates to novel human coding sequences (ABO99268-  
 ABO99608) and proteins (ABP6482-ABP65022). The sequences are useful in  
 therapeutic, diagnostic and research methods. The polynucleotides may be  
 used in the field of molecular biology as hybridisation probes, primers  
 for PCR, for chromosome and gene mapping, for the recombinant production  
 of protein, or in generation of anti-sense DNA or RNA. The  
 polynucleotides are useful in diagnostics as expressed sequence tags  
 (ESTs) for identifying expressed genes or for physical mapping of the  
 human genome. The proteins may be used as molecular weight markers, or as  
 nutritional sources or supplements. The proteins may be used to maintain  
 and expand cell population in a totipotential or pluripotential state  
 useful for re-engineering damaged or diseased tissues, transplantation,  
 manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
 polynucleotides and proteins are useful for preventing, treating or  
 ameliorating disorders involving aberrant protein expression or  
 biological activity, e.g. haematopoietic disorders, central/peripheral  
 nervous system diseases, mechanical and traumatic disorders, non-healing  
 wounds, immune deficiencies and disorders, infectious diseases caused by  
 viral, bacterial or fungal infection, autoimmune disorders, allergic  
 reactions and conditions, coagulation disorders, or cancer. The  
 polynucleotide sequences of the invention were assembled from ESTs  
 isolated mainly by sequencing by hybridisation, and in some cases,  
 sequences obtained from one or more public databases. Note: The sequence  
 data for this patent did not form part of the printed specification, but  
 was obtained in electronic format directly from WIPO at  
 ftp.wipo.int/pub/published\_pct\_sequences

Sequence 210 AA;

Query Match 68.3%; Score 923.5; DB 5; Length 210;  
 Best Local Similarity 87.9%; Pred. No. 2.1e-86;  
 Matches 182; Conservative 3; Mismatches 15; Indels 7; Gaps 1;

1 MGSLPSRRKSLPSLSSSVQCGPVTWEAERSKATAVALGSPFAGGPAELSLRLGPLET 60  
 1 MGSLPSRRKSLPSLSSSVQCGPVTWEAERSKATAVALGSPFAGGPAELSLRLGPLET 60

61 IVSEGDGNTVLSVSGREYNTPSVHVGVKSHGWLVEGLSRKABELLLPNCGGAFLL 120  
 61 IVSEGDGNTVLSVSGREYNTPSVHVGVKSHGWLVEGLSRKABELLLPNCGGAFLL 120

QY 121 RESQTRRGYSLSVRLSRPASWDRIHYRIHCLDNGWLYISFRLTFPSLQALVDHYSE-- 178  
 Db 121 RESQTRRGYSLSVRLSRPASWDRIHYRIHCLDNGWLYISFRLTFPSLQALVDHYSEGW 180

QY 179 -----LADDICCLLKEPCVLRAGPLP 200  
 Db 181 PAPWQGYTPTCDCAEDTTQLERAGQLP 207

RESULT 9  
 ABP64107  
 ID ABP64107 standard; protein; 159 AA.  
 XX ABP64107;  
 XX  
 DT 04-NOV-2002 (first entry)  
 XX  
 DE Human ORF477.  
 XX  
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FN US2002082206-A1.  
 XX  
 PD 27-JUN-2002.  
 XX  
 PF 30-MAY-2001; 2001US-00867550.  
 XX  
 PR 30-MAY-2000; 2000US-0208427P.  
 PA (LEAC/) LEACH M D.  
 PA (MEHR/) MEHRABAN F.  
 PA (CONL/) CONLEY P B.  
 PA (TOPP/) TOPPER J N.  
 PA (LAWD/) LAW D.  
 XX  
 PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX  
 DR WPI; 2002-626554/67.  
 DR N-PSDB; ABO98670.  
 XX  
 PT New polypeptide designated ORFX are present in human atherogenic cells  
 and are useful to prevent and treat ORFX-associated disorders including  
 cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 XX  
 PS Claim 10; SEQ ID NO 954; 78pp; English.  
 XX  
 CC The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABO98194-ABO99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?DocID=2002082206  
 XX  
 CC Sequence 159 AA;

Query Match 61.0%; Score 826; DB 5; Length 159;  
 Best Local Similarity 99.4%; Pred. No. 1.5e-76;  
 Matches 159; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSLSRKRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGEPLT 60  
Db 1 MGSLSRKRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGEPLT 60  
QY 61 IVSEDDGWTWLVSEVSGREYNIPSVHGVKSHGLYEGLSREKAEELLLLPGNPGGAFLI 120  
Db 61 IVSEDDGWTWLVSEVSGREYNIPSVHGVKSHGLYEGLSREKAEELLLLPGNPGGAFLI 120  
QY 121 RESQTRGSSLSVRLSRPASWDRIRHYRTHCLDNGWLY 159  
Db 121 RESQTRGSSLSVRLSRPASWDRIRHYRTHCLDNGWLY 159

RESULT 10  
ABP64588  
ID ABP64588 standard; protein; 113 AA.  
XX  
AC ABP64588;  
XX  
DT 04-NOV-2002 (first entry)  
XX  
DE Human ORF958.  
XX  
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnery;  
KW Antinflammatory; Gene therapy; human; ORFX; atherogenic; platelet;  
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
KW cancer; cardiovascular disease; allergy; autoimmune disease;  
KW wound healing; blood coagulation disorder; inflammatory disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2002082206-A1.  
XX  
PD 27-JUN-2002.  
XX  
PF 30-MAY-2001; 2001US-00867550.  
XX  
PR 30-MAY-2000; 2000US-0208427P.  
XX  
PA (LEAC/) LEACH M D.  
PA (MEHR/) MEHRABAN F.  
PA (CONL/) CONLEY P B.  
PA (TOPP/) TOPPER J N.  
PA (LAWD/) LAW D.  
XX  
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
XX  
DR WPI; 2002-626554/67.  
XX  
DR N-PSDB; ABQ99151.  
XX  
PT New polypeptide designated ORFX are present in human atherogenic cells  
PT and are useful to prevent and treat ORFX-associated disorders including  
PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
PT inflammatory disease.  
XX  
PS Claim 10; SEQ ID NO 1916; 78pp; English.  
XX  
CC The present invention relates to novel human ORFX polypeptides and their  
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
CC were discovered in human atherogenic cells, in particular in platelets  
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
CC many other tissues as well. Atherogenic cells are cells which have the  
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
CC nucleic acids are useful for treating or preventing a pathological  
CC condition associated with an ORFX-associated disorder, e.g. cancer,  
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
CC coagulation disorders or inflammatory disorders. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from the USPTO web site at  
CC seqdata.uspto.gov/sequence.html?DocID=2002082206  
XX  
SQ Sequence 113 AA;

Query Match 43.3%; Score 586; DB 5; Length 113;  
Best Local Similarity 100.0%; Pred. No. 4.7e-52;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 150 IHLCDNGWLYISPRITPSPLOALVDHYSELADICLLKEPCVLQACPLPGKDIPLPVT 209  
Db 2 IHLCDNGWLYISPRITPSPLOALVDHYSELADICLLKEPCVLQACPLPGKDIPLPVT 61  
QY 210 VQRTPLNWKELDSSLFSEAAATGSESLSEGRLSFYISLNDEAVSLDDA 261  
Db 62 VQRTPLNWKELDSSLFSEAAATGSESLSEGRLSFYISLNDEAVSLDDA 113

RESULT 11  
ABP52188  
ID ABP52188 standard; protein; 276 AA.  
XX  
AC ABP52188;  
XX  
DT 15-OCT-2002 (first entry)  
XX  
DE Human Src-like adapter protein (SLAP).  
XX  
KW Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;  
KW SLAP; inhibitor; anti-inflammatory; immunosuppressive; anti-HIV;  
KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;  
KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
KW chronic inflammatory disorder; autoimmune disorder; transplant rejection.  
XX  
OS Homo sapiens.  
XX  
PN WO200255707-A2.  
XX  
PD 18-JUL-2002.  
XX  
PF 10-JAN-2002; 2002WO-US000718.  
XX  
PR 10-JAN-2001; 2001US-0260953P.  
XX  
PA (RIGE-) RIGEL PHARM INC.  
XX  
PI Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
XX  
DR WPI; 2002-575432/61.  
XX  
PT New src-like inhibitory molecule protein, useful for treating  
PT immunodeficiency disorders and inflammatory disorders, comprises N-  
PT terminal myristylation sequence, SH2 domain and/or SH3 domain.  
XX  
PS Example 2; Fig 2B; 91pp; English.  
XX  
CC The present invention describes the human Src-like inhibitory molecule  
CC (SLIM) protein (I). The present invention describes a SLIM protein  
CC comprising an N-terminal myristylation sequence, an N-terminal SH2  
CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising  
CC an N-terminal myristylation sequence and an N-terminal SH2 domain which  
CC is unable to bind to Cbl. (I) has anti-inflammatory, immunosuppressive and  
CC anti-HIV activities, and can be used as a modulator of lymphocyte  
CC activation, and of ubiquitination of a Cbl target protein, and in gene  
CC therapy. (I) is useful for screening a bioactive agent capable of binding  
CC to SLIM. (I) is also useful for screening a bioactive agent capable of  
CC modulating SLIM binding. (I) or its fragments is useful in the study or  
CC in the treatment of conditions which involve this function or  
CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or  
CC prevent SLIM associated disorders. (I) or the polynucleotide encoding it  
CC (II) is useful for modulating leukocyte and/or platelet activation, for  
CC modulating antigen receptor-induced signaling and activation in  
CC leukocyte and/or platelets and for modulating antigen receptor-induced  
CC signaling and activation in lymphocytes and/or mast cells. (I) or (II)  
CC is also useful for modulating the basal activity of lymphocytes. (I) or  
CC (II) is useful in the treatment of immunodeficiency disorders, such as

CC acquired immunodeficiency syndrome (AIDS), for the prevention and  
 CC treatment of acute inflammatory disorders, chronic inflammatory  
 CC disorders, autoimmune disorder and transplant rejection. The present  
 CC sequence represents the human Src-like adapter protein (SLAP), which is  
 CC given in comparison with the human SLIM protein in an example from the  
 CC present invention  
 XX  
 SQ Sequence 276 AA;  
 Query Match 35.6%; Score 481.5; DB 5; Length 276;  
 Best Local Similarity 40.3%; Pred. No. 1.1e-40;  
 Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;  
 QY 9 KSLPSPSLSSVQGGPVTMEAESKATAVALGSPGAGPAELSLRLGEPITIVSEGDW 68  
 Db 6 KSTPAPA-----ERPLNPEGLSDFLAVLSDYSPDISPPIFRGKELRVISDEGW 58  
 QY 69 WTVLSEVSGREYNIPSVHGVKSHGWLVEG-SREKAEELLLLPGNPGGAFILRESQTRRG 128  
 Db 59 WKAISLSTGRESYIPGICVARVYHGWLFEGGLDRKAEELLQLPDKVGSFMRSETKKG 118  
 QY 129 SYSLSVRLSRPASWDRIRHYRHCLDNGWLVTISPELTSPSLQALVDHYSELADICCLLK 188  
 Db 119 FYSLSVR-----HRQVKHYRIFRLPNWYIISPLTFCLEDLVNHVSEVADGLCCVLT 172  
 QY 189 EPCVLQAGPLFGKDIPLPVTQRTPLNWKELDSSLLFSEAATG-----EESLLSEGL 241  
 Db 173 TPCLTQSTAAPAVRASSSPVTLRQKTVDMRRVSR---LQEDPEGTENPLGVDESLSFVGL 229  
 QY 242 RESLSFYISLNDE 254  
 Db 230 RESIASYLSLTSE 242  
 RESULT 12  
 ID ABU79103 standard; protein; 276 AA.  
 XX  
 AC ABU79103;  
 XX  
 DT 18-JUN-2003 (first entry)  
 DE Src-like adaptor protein, SLAP.  
 XX  
 KW Superantigen; SAG; staphylococcal enterotoxin; tumour; cancer; apoptosis;  
 KW gene therapy; mammalian cell receptor; cytostatic;  
 KW tumour associated lipid; energy; T cell; antigen presenting cell; APC;  
 KW tumouricidal immunocyte; antitumour.  
 XX  
 OS Unidentified.  
 XX  
 PN US2002177551-A1.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 30-MAY-2001; 2001US-00870759.  
 XX  
 PR 31-MAY-2000; 2000US-0208128P.  
 XX  
 PA (TERM/) TERMAN D S.  
 XX  
 PI Terman DS;  
 XX  
 DR WPI; 2003-361759/34.  
 XX  
 PT A mammalian cell receptor, useful in the treatment of cancer, by binding  
 PT to tumor associated lipids where the binding induces energy or apoptosis  
 PT in T cells and antigen presenting cells.  
 XX  
 PS Disclosure; Page; 167pp; English.  
 XX  
 CC The invention relates to a mammalian cell receptor, useful in the  
 CC treatment of cancer, which binds to tumour associated lipids and induces

CC energy or apoptosis in the T cells and antigen presenting cells (APCs).  
 CC Also included are a mammalian cell useful in the treatment of cancer  
 CC where the receptor which binds tumour associated lipids and induces  
 CC cellular inactivation or death is deleted or functionally deactivated,  
 CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal  
 CC (by allowing tumour associated lipids to contact immunocytes in which  
 CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,  
 CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,  
 CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or  
 CC deleted), a construct useful in the treatment of cancer comprising a  
 CC superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell  
 CC useful in the treatment of cancer (where an adaptor protein which  
 CC inhibits T cell activation by tumour associated antigens is deleted or  
 CC functionally deactivated), a composition useful in the treatment of  
 CC cancer (comprising a lipid raft conjugated to a superantigen), producing  
 CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by  
 CC allowing tumour associated lipids to contact immunocytes, in which  
 CC receptors for the lipids are inactivated or deleted to produce a  
 CC tumouricidal immunocyte population, and administering the tumouricidal  
 CC activated immunocytes to the host), producing (M3) a tumouricidal APC  
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to  
 CC contact APCs, in which receptors for the tumour associated lipids are  
 CC inactivated or deleted to produce a tumouricidally activated population,  
 CC and administering APCs to the host), producing a tumouricidal T cell  
 CC population ex vivo in a mammal (by allowing a tumour associated lipids to  
 CC contact T cells, in which adaptor proteins, which inhibit T cell  
 CC activation by tumour associated antigens, are deleted or functionally  
 CC deactivated to produce a tumouricidal population of T cells, and  
 CC administering the tumouricidally activated T cells to the host, or  
 CC allowing a superantigen-lipid raft to contact T cells ex vivo, and  
 CC administering the tumouricidally activated T cells to the host), treating  
 CC (M5) cancer in a mammal (by administering a lipid binding molecule which  
 CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)  
 CC a tumouricidal T cell population in vivo in a mammal (by allowing a  
 CC tumour associated antigen to contact immunocytes in which adaptor  
 CC proteins which inhibit T cell activation by tumour associated antigens  
 CC are deleted or functionally deactivated) and producing (M7) a  
 CC tumouricidal T cell population ex vivo in a mammal comprising allowing a  
 CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The  
 CC receptors, methods and compositions are useful for treating cancers and  
 CC tumours. Bacterial superantigens are co-administered or administered as  
 CC fusion constructs with anti-tumour proteins or motifs. The present  
 CC sequence represents a tumour antigen or a motif identifying a tumour  
 CC antigen, which can be functionally deactivated in the method of the  
 CC invention. Note: the sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format from the  
 CC US patent office website at  
 CC "seqdata.uspto.gov/sequence.html?DocID=20020177551"  
 XX  
 SQ Sequence 276 AA;

Query Match 35.6%; Score 481.5; DB 6; Length 276;  
 Best Local Similarity 40.3%; Pred. No. 1.1e-40;  
 Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;  
 QY 9 KSLPSPSLSSVQGGPVTMEAESKATAVALGSPGAGPAELSLRLGEPITIVSEGDW 68  
 Db 6 KSTPAPA-----ERPLNPEGLSDFLAVLSDYSPDISPPIFRGKELRVISDEGW 58  
 QY 69 WTVLSEVSGREYNIPSVHGVKSHGWLVEG-SREKAEELLLLPGNPGGAFILRESQTRRG 128  
 Db 59 WKAISLSTGRESYIPGICVARVYHGWLFEGGLDRKAEELLQLPDKVGSFMRSETKKG 118  
 QY 129 SYSLSVRLSRPASWDRIRHYRHCLDNGWLVTISPELTSPSLQALVDHYSELADICCLLK 188  
 Db 119 FYSLSVR-----HRQVKHYRIFRLPNWYIISPLTFCLEDLVNHVSEVADGLCCVLT 172  
 QY 189 EPCVLQAGPLFGKDIPLPVTQRTPLNWKELDSSLLFSEAATG-----EESLLSEGL 241  
 Db 173 TPCLTQSTAAPAVRASSSPVTLRQKTVDMRRVSR---LQEDPEGTENPLGVDESLSFVGL 229  
 QY 242 RESLSFYISLNDE 254



Db 230 RESIASYLSLTSE 242

RESULT 13

ABR59698

ID ABR59698 standard; protein; 276 AA.

AC ABR59698;

XX

DT 25-JUL-2003 (first entry)

XX

DE Human Src-like adaptor.

XX

XX Human; T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; asthma;

KW immunosuppressive; antiaesthatic; anti-allergic; anti-inflammatory;

KW lymphocyte activation; lymphocyte migration; cytokine production;

KW cell surface marker expression; antibody production; apoptosis; allergy;

KW antibody proliferation; antibody differentiation; hypersensitivity;

KW graft versus host disease; inflammation; Src-like-adaptor.

XX

OS Homo sapiens.

XX

XX WO2003029277-A2.

PN

XX

PD 10-APR-2003.

XX

PF 02-OCT-2002; 2002WO-US031618.

XX

PR 03-OCT-2001; 2001US-0327212P.

XX

PA (RIGB-) RIGEL PHARM INC.

XX

PI Chu P, Li C, Liao XC, Masuda E, Pardo J, Zhao H;

PI

XX WPI; 2003-363276/34.

DR

DR N-PSDB; ACC81091.

XX

XX

PT Identifying a compound that modulates T lymphocyte activation, useful for

PT monitoring changes in cell surface marker expression, comprises

PT contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with

PT a compound.

PT

XX Disclosure; Page 72; 126pp; English.

PS

XX

XX The invention relates to a novel method for identifying a compound that

CC modulates T lymphocyte activation. The method comprises contacting a T

CC cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,

CC where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic

CC acid that hybridises to a nucleic acid encoding a polypeptide having a

CC sequence selected from two 606-amino acid sequence and a 415-amino acid

CC sequence given in the specification. The method of the invention has

CC immunosuppressive, antiaesthatic, anti-allergic, and anti-inflammatory

CC activity. The method is useful for identifying compounds that modulate

CC lymphocyte activation and migration, and for monitoring changes in cell

CC surface marker expression, cytokine production, antibody production,

CC proliferation and differentiation, and apoptosis, using either cell lines

CC or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as

CC drug targets for compounds that suppress or activate lymphocyte

CC activation and migration, e.g. for the treatment of diseases in which

CC modulation of the immune response is desired such as delayed type

CC hypersensitivity reactions, asthma, allergies, graft versus host disease,

CC and acute and chronic inflammation. Modulators of lymphocyte activation

CC are useful for treating disorders related T and B cell activation and

CC migration. The present sequence is used in the exemplification of the

CC invention

XX

XX Sequence 276 AA;

XX

Query Match 35.6%; Score 481.5; DB 6; Length 276;

Best Local Similarity 40.3%; Pred. No. 1.1e-40;

Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;

9 KSLPSPSLSSVQGGQPVTVWEABRSKATVALGSPFAGGPAELSLRLGEPLTVSEDDGW 68

Db 6 KSTPAPA-----ERLPNPEGLSDFLVLSYSPDISPPIFRGKLRVISDEGGW 58

QY 69 WTVLSEVSGREYNIPSVHGVKSHGWLVEGLSREKABELLLPGNPGGAFILRESQTRRG 128

Db 59 WKAISSLTGRESYIPGICVARVYHGLFEGLRDKABELLQLPDTKVGSFMRSETKKG 118

QY 129 SYSLSVRLSRPASWDRIHRIHCLDNGWLYISPLTFPSLQALVDHYSELADICLLK 188

Db 119 FYLSLSVR-----HRQVKHYRIFELPNNWYIISPLTFQCLELVDVHYSEVADGLCCVLT 172

QY 189 EPCVLQRAGPLPGKDIPLPVTQVORTPLNWKELSSLLFSBAATG-----EESLLSEGL 241

Db 173 TPCLTQTSTAAPAVRASSPVTLRKQTVDMRVSRLQEDPEGTENPLGVDESLFSYGL 229

QY 242 RESLSFYISLND 254

Db 230 RESIASYLSLTSE 242

RESULT 14

AAU31072

ID AAU31072 standard; protein; 315 AA.

XX

AC AAU31072;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secreted protein #1563.

XX

KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX

OS Homo sapiens.

XX

XX WO200179449-A2.

PN

XX

PD 25-OCT-2001.

XX

PF 16-APR-2001; 2001WO-US008656.

XX

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-611725/70.

XX

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

PT vaccination, testing and therapy.

PT

XX Claim 20; Page 399; 765pp; English.

XX

CC The invention relates to novel human secreted polypeptides. The

CC polypeptides and antibodies to the polypeptides are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising the

CC nucleic acids encoding the polypeptides and cells genetically engineered

CC to express them are also useful for producing the proteins. The proteins

CC are useful in genetic vaccination, testing and therapy, and can be used

CC as nutritional supplements. They may be used to increase stem cell

CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon

CC and/or nerve tissue growth or regeneration; immune suppression and/or

CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.

CC AAU29310-AAU33304 represent the amino acid sequences of novel human

CC secreted proteins of the invention



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:19:42 ; Search time 45 Seconds  
(without alignments)  
557.911 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLSRRKSLSPSLSSSV.....RESLSFYSLNDEAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	493.5	36.5	281	2 A57152	src-like adaptor p
2	370.5	27.4	512	1 TVHULY	protein-tyrosine k
3	367.5	27.2	512	1 I56160	protein-tyrosine k
4	367.5	27.2	512	1 A39719	protein-tyrosine k
5	360.5	26.6	505	1 TVHUHC	protein-tyrosine k
6	353.5	26.1	503	1 JQ1321	protein-tyrosine k
7	352.5	26.1	505	2 I37206	protein-tyrosine k
8	352	26.0	503	1 TVMSHC	protein-tyrosine k
9	344.5	25.5	499	1 A40092	protein-tyrosine k
10	340	25.1	509	1 OKHULK	protein-tyrosine k
11	338.5	25.0	509	1 I48945	protein-tyrosine k
12	333	24.6	507	1 A39939	protein-tyrosine k
13	325	24.0	539	2 B49114	protein-tyrosine k
14	320	23.7	528	1 TVFVG9	protein-tyrosine k
15	319.5	23.6	537	1 A45501	protein-tyrosine k
16	317	23.4	541	1 TVCHYS	protein-tyrosine k
17	315.5	23.3	537	1 TVHUSY	protein-tyrosine k
18	315.5	23.3	543	1 TVHUSY	protein-tyrosine k
19	312.5	23.1	529	1 TVHUFR	protein-tyrosine k
20	311	23.0	544	2 I51593	protein-tyrosine k
21	310.5	22.9	537	1 A43806	protein-tyrosine k
22	309	22.8	534	1 A44991	protein-tyrosine k
23	307	22.7	517	2 S24547	protein-tyrosine k
24	306.5	22.7	541	2 S31645	protein-tyrosine k
25	305.5	22.6	537	2 I51592	protein-tyrosine k
26	301	22.2	517	2 A43807	protein-tyrosine k
27	301	22.2	536	2 S33569	protein-tyrosine k
28	297.5	22.0	534	1 S33568	protein-tyrosine k
29	295.5	21.8	542	2 A49114	protein-tyrosine k

30	290	21.4	506	1 S24553	protein-tyrosine k
31	286.5	21.2	557	1 TVFVS2	protein-tyrosine k
32	286.5	21.2	587	1 TVFVPR	protein-tyrosine k
33	284.5	21.0	533	1 TVCHS	protein-tyrosine k
34	284.5	21.0	568	1 TVFV51	protein-tyrosine k
35	279.5	20.7	526	1 TVFV60	protein-tyrosine k
36	276	20.4	509	1 TVHAST	protein-tyrosine k
37	275	20.3	546	2 S52314	protein-tyrosine k
38	274.5	20.3	532	1 B34104	protein-tyrosine k
39	273.5	20.2	526	1 TVFVR	protein-tyrosine k
40	273.5	20.2	526	2 S15582	protein-tyrosine k
41	273.5	20.2	542	1 TVHUSC	protein-tyrosine k
42	273	20.2	545	2 S52313	protein-tyrosine k
43	272	20.1	541	1 A43610	protein-tyrosine k
44	271.5	20.1	532	1 A34104	protein-tyrosine k
45	269.5	19.9	526	2 S26420	protein-tyrosine k

ALIGNMENTS

RESULT 1

A57152

src-like adaptor protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 08-Dec-1995 #sequence\_revision 08-Dec-1995 #text\_change 12-Feb-1999

C:Accession: A57152

R:pandey, A.; Duan, H.; Dixit, V.M.

J. Biol. Chem. 270, 19201-19204, 1995

A>Title: Characterization of a novel Src-like adapter protein that associates with the Bcr

A:Reference number: A57152; MUID:95370243; PMID:7543898

A:Accession: A57152

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-261 <PAN>

A:Cross-references: GB:U29056

C:Superfamily: SH3 homology; SH2 homology

F:29-77/Domain: SH3 homology <SH3>

F:84-175/Domain: SH2 homology <SH2>

Query Match 36.5%; Score 493.5; DB 2; Length 281;  
Best Local Similarity 43.6%; Pred. No. 2.1e-33;  
Matches 115; Conservative 36; Mismatches 82; Indels 29; Gaps 7;

QY	9	KSLPSPS---	LSSSVQCGPVTMEASRATKATVALGFPAGPAELSLRLGELPTIVSE	65
Db	6	KSTSPSPSERPLSSS-	-----EGLESDFLAV-LTDYPSSDISPPIFRGRGKLRVISDE	55
QY	66	GDWMTVLSEVSGREYNTIPSVHVGKVSCHWLYEGLSRKAEELLLPGNPGGAPLIRE	SQT 125	
Db	56	GGWKAISLSTGRESYIPGICVARVHGWLFEGLRDKAEELQLPPTKIGSPMIRE	SET 115	
QY	126	RRGSYSLSVRLSRPASWDRIHRIHCLDNGWLYISPRITFPSPQLQALVDHYSELAD	ICC 185	
Db	116	KKGYSLSVR-----HRQVKHYRIPELPNNWYISPRITFPQCLEDLVTHYSE	ADGLCC 169	
QY	186	LLKEPCVQLR-----AGLPCKDIPLVTVQRTPLNWKELDSSLLFSEATG	-----BESL 236	
Db	170	VLTPFCLAQIPAPTSPCTSPGSPVTLRQKTFDWRVSRQEGSEGAENPLRVDE	SL 229	
QY	237	LSEGLRESLSFYSL-NDEAVSLD	259	
Db	230	PSYGLRESIASYLSLTGDDSSSP	253	

RESULT 2

TVHULY

protein-tyrosine kinase (EC 2.7.1.112) lyn, splice form A - human

N:Contains: protein-tyrosine kinase lyn, splice form B

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 18-Feb-2000

C:Accession: A26719; D38268; PH0949; I53715

R:Yamanashi, Y.; Fukushima, S.I.; Semba, K.; Sukegawa, J.; Miyajima, N.; Matsubara, K.;

Mol. Cell. Biol. 7, 237-243, 1987  
A;Title: The yes-related cellular gene lyn encodes a possible tyrosine kinase similar to  
A;Reference number: A26719; MUID:87172710; PMID:3561390  
A;Accession: A26719  
A;Molecule type: mRNA  
A;Residues: 1-512 <VAM>  
A;Cross-references: GB:M16038; NID:g187268; PIDN:AA859540.1; PID:g307144  
R;Paranen, J.; Mäkelä, T.P.; Allitalo, R.; Lehtvaeslahti, H.; Allitalo, K.  
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990  
A;Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.  
A;Reference number: A39268; MUID:91062389; PMID:2247464  
A;Accession: D38268  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 369-424 <PAR>  
R;Bieleke, W.; Ziemleki, A.; Kappos, L.; Miescher, G.C.  
Biochem. Biophys. Res. Commun. 186, 1403-1409, 1992  
A;Title: Expression of the B cell-associated tyrosine kinase gene lyn in primary neurobl  
A;Reference number: PH0949; MUID:92378604; PMID:1510669  
A;Accession: PH0949  
A;Molecule type: mRNA  
A;Residues: 369-424 <BIE>  
A;Experimental source: neuroblastoma SK-IN cell  
R;Rider, L.G.; Raben, N.; Miller, L.; Jelsema, C.  
Gene 138, 219-222, 1994  
A;Title: The cDNAs encoding two forms of the LYN protein tyrosine kinase are expressed i  
A;Reference number: I53715; MUID:94171041; PMID:8125304  
A;Accession: I53715  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-24,46-512 <RID>  
A;Cross-references: GB:M79321; NID:g187270; PIDN:AA850019.1; PID:g187271  
A;Experimental source: splice form B  
C;Genetics:  
A;Gene: GDB:LYN  
A;Cross-references: GDB:120159; OMIM:165120  
A;Map position: 8q13-8qter  
C;Function:  
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprote  
tyrosine-specific protein kinase  
F;2-512/Product: protein-tyrosine kinase  
F;2-24,46-512/Product: protein-tyrosine kinase lyn, splice form A #status predicted <MATA>  
F;70-118/Domain: SH3 homology <SH3>  
F;129-226/Domain: SH2 homology <SH2>  
F;245-504/Domain: protein kinase homology <KIN>  
F;253-261/Region: protein kinase ATP-binding motif  
F;3/Binding site: palmitate (Cys) (covalent) #status predicted  
F;275/Active site: Lys #status predicted  
F;397,508/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 27.4%; Score 370.5; DB 1; Length 512;  
Best Local Similarity 39.8%; Pred. No. 6.9e-23;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
QY 6 SRKSLPSLSVSSVGGPVTWEAERSKATAVALGSPAGGPAELSLRLGPELTVSSED 65  
DB 38 SNKQRPVPE-SQLLPQGRFQKDPPEEQGDIVVLYPYDGIHPDDLSFKKGKMKVLEEH 96  
QY 66 GDMWTVLSEVSGREYNIPSVHVGK-----SHGWLVEGLSREKAEELLLPGNPGGAFLIR 121  
DB 97 GEWKAJSLTKKEGFIIPSNVAKNTLETEWFFKDIKDAERQLLAPGNSAGAFILIR 156  
QY 122 ESOTRGSVLSVLRSPASWRIRHYRHCLDNGWLVIYSPRLTPPSLQALVDHYSELAD 181  
DB 157 ESETLKGSPSLVRDPPVHGDVIRHYKIRSLDNGYIYSPRITPPCISDMIKHYKQAD 216  
QY 182 DICLLKPCVLRQAGPLPK 202  
DB 217 GLCRLKAKI-----SPKPK 233

RESULT 3  
I56160  
N;Contains: protein-tyrosine kinase lyn, splice form A - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 18-Feb-2000 #sequence revision 18-Feb-2000 #text\_change 18-Feb-2000  
C;Accession: I56160; I67811; I67812  
R;Minoguchi, K.; Nishikata, H.; Siraganian, R.P.  
J. Immunol. 150, 222, 1993  
A;Title: Bacterially expressed rat p56lyn binds several proteins in rat basophilic leuke  
A;Reference number: I56160  
A;Accession: I56160  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-512 <MIN>  
A;Cross-references: GB:I14951; NID:g294582; PIDN:AAA41549.1; PID:g294583  
R;Rider, L.G.; Raben, N.; Miller, L.; Jelsema, C.  
Gene 138, 219-222, 1994  
A;Title: The cDNAs encoding two forms of the LYN protein tyrosine kinase are expressed i  
A;Reference number: I53715; MUID:94171041; PMID:8125304  
A;Accession: I67811  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-230,'L',232-307,'A',309-418,'Y',420-512 <RID1>  
A;Cross-references: GB:I14782; NID:g294578; PIDN:AAA20944.1; PID:g294579  
A;Note: in Genbank entry RATLYNATYR, release 116.0, PIDN:AAA20944.1, the source is desig  
A;Accession: I67812  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-24,46-230,'L',232-307,'A',309-418,'Y',420-512 <RID2>  
A;Cross-references: GB:I14823; NID:g294580; PIDN:AAA20945.1; PID:g294581  
A;Note: in Genbank entry RATLYNATYR, release 116.0, PIDN:AAA20945.1, the source is desig  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprote  
F;2-512/Product: protein-tyrosine kinase lyn, splice form A #status predicted <MATA>  
F;2-24,46-512/Product: protein-tyrosine kinase lyn, splice form B #status predicted <MAT  
F;70-118/Domain: SH3 homology <SH3>  
F;129-226/Domain: SH2 homology <SH2>  
F;245-504/Domain: protein kinase homology <KIN>  
F;253-261/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;275/Active site: Lys #status predicted  
F;397,508/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 27.2%; Score 367.5; DB 1; Length 512;  
Best Local Similarity 40.3%; Pred. No. 1.2e-22;  
Matches 81; Conservative 33; Mismatches 78; Indels 9; Gaps 3;  
QY 6 SRKSLPSLSVSSVGGPVTWEAERSKATAVALGSPAGGPAELSLRLGPELTVSSED 65  
DB 38 SNKQRPVPE-SQLLPQGRFQKDPPEEQGDIVVLYPYDGIHPDDLSFKKGKMKVLEEH 96  
QY 66 GDMWTVLSEVSGREYNIPSVHVGK-----SHGWLVEGLSREKAEELLLPGNPGGAFLIR 121  
DB 97 GEWKAJSLTKKEGFIIPSNVAKNTLETEWFFKDIKDAERQLLAPGNSAGAFILIR 156  
QY 122 ESOTRGSVLSVLRSPASWRIRHYRHCLDNGWLVIYSPRLTPPSLQALVDHYSELAD 181  
DB 157 ESETLKGSPSLVRDPPVHGDVIRHYKIRSLDNGYIYSPRITPPCISDMIKHYKQSD 216  
QY 182 DICLLKPCVLRQAGPLPK 202  
DB 217 GLCRLKAKI-----SPKPK 233

RESULT 4  
A39719  
N;Contains: protein-tyrosine kinase (EC 2.7.1.112) lyn, long splice form - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 18-Feb-2000 #sequence revision 18-Feb-2000 #text\_change 03-Mar-2000  
C;Accession: A39719; B39719; A39750; B39750

R;Stanley, E.; Ralph, S.; McEwen, S.; Boulet, I.; Holtzman, D.A.; Lock, P.; Dunn, A.R.  
Mol. Cell. Biol. 11, 3399-3406, 1991  
A;Title: Alternatively spliced murine lyn mRNAs encode distinct proteins.  
A;Reference number: A39719; MUID:91260688; PMID:1710766  
A;Accession: A39719  
A;Molecule type: mRNA  
A;Residues: 1-512 <STAL>  
A;Cross-references: GB:M64608; NID:g198938; PIDN:AAA39470.1; PID:g198939  
A;Accession: B39719  
A;Molecule type: mRNA  
A;Residues: 1-24,46-512 <STA2>  
A;Cross-references: GB:M64608  
R;Yi, T.; Bolen, J.B.; Ihle, J.N.  
Mol. Cell. Biol. 11, 2391-2398, 1991  
A;Title: Hematopoietic cells express two forms of lyn kinase differing by 21 amino acids  
A;Reference number: A39750; MUID:91203857; PMID:2017160  
A;Accession: A39750  
A;Molecule type: mRNA  
A;Residues: 1-76,'F',78-160,'I',162-278,'L',280-390,'I',392-424,'D',426-512 <YI1>  
A;Cross-references: GB:M57696; NID:g198940; PIDN:AAA39471.1; PID:g198941  
A;Accession: B39750  
A;Molecule type: mRNA  
A;Residues: 1-24,46-76,'B',78-160,'I',162-278,'L',280-390,'I',392-424,'D',426-512 <YI2>  
A;Cross-references: GB:M57697; NID:g198942; PIDN:AAA39472.1; PID:g198943  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprote  
F;1-512/Product: protein-tyrosine kinase lyn, long splice form #status predicted <MATL>  
F;1-24,46-512/Product: protein-tyrosine kinase lyn, short splice form #status predicted  
F;70-118/Domain: SH3 homology <SH3>  
F;129-226/Domain: SH2 homology <SH2>  
F;243-504/Domain: protein kinase homology <KIN>  
F;253-261/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;275/Active site: Lys #status predicted  
F;397,508/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 27.2%; Score 367.5; DB 1; Length 512;  
Best Local Similarity 40.3%; Pred. No. 1.2e-22;  
Matches 81; Conservative 33; Mismatches 78; Indels 9; Gaps 3;

QY 6 SRRKSLPSPS-SSSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPLTIVSED 65  
Db 38 SNKQRPVPEF-HLLPGQRFQTKPEEGDIVVALYFDGHPDDLSFKKGGKMKVLEH 96  
QY 66 GDWTVLSEVSGREYNIPSVHGVK- ---SHGWLYEGLSRKAEELLILLPGNPGAPLIR 121  
Db 97 GEWKAKESSKREGFIPSNVAKVNTLETBEWFFKDIITRDAERQLLAPGNSAGAPLIR 156  
QY 122 ESQTRRGYSLSVRLSRPASWDRIHRHYRIHCLDNGWLYISPRITFPPSLQALVDHYSLEAD 181  
Db 157 ESETLKGSFSLVSDYDPMHGDVVKHYKIRSLDNGGYIISPRITFPCISDMIKHYQKSD 216  
QY 182 DICLLKEPCVLRAGPLPGK 202  
Db 217 GLCRRLEXACI- ---SPKPK 233

RESULT 5  
TVHHC  
protein-tyrosine kinase (EC 2.7.1.112) hck - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1989 #sequence revision 10-Nov-1995 #text\_change 11-Jun-1999  
C;Accession: A27811; A27812; JCI149; C38268; S31103  
R;Quintrell, N.; Lebo, R.; Varmus, H.; Bishop, J.M.; Pettenati, M.J.; Le Beau, M.M.; Dia  
Mol. Cell. Biol. 7, 2267-2275, 1987  
A;Title: Identification of a human gene (HCK) that encodes a protein-tyrosine kinase and  
A;Reference number: A27811; MUID:87257942; PMID:3496523  
A;Accession: A27811  
A;Molecule type: mRNA  
A;Residues: 1-505 <QUT>  
A;Cross-references: GB:M16591  
A;Note: the codon given for 3-Cys (TCG) is inconsistent with the authors' translation  
R;Ziegler, S.F.; Marth, J.D.; Lewis, D.B.; Perlmutter, R.M.

Mol. Cell. Biol. 7, 2276-2285, 1987  
A;Title: Novel protein-tyrosine kinase gene (hck) preferentially expressed in cells of he  
A;Reference number: A27812; MUID:87257943; PMID:3453117  
A;Accession: A27812  
A;Molecule type: mRNA  
A;Residues: 1-505 <ZIE>  
A;Cross-references: GB:M16592; NID:g183913; PIDN:AAA52644.1; PID:g306833  
R;Hradetzky, D.; Strebhardt, K.; Ruebsamen-Waligmann, H.  
Gene 113, 275-280, 1992  
A;Title: The genomic locus of the human hemopoietic-specific cell protein tyrosine kinase  
A;Reference number: JCI149; MUID:92241680; PMID:1572549  
A;Accession: JCI149  
A;Molecule type: DNA  
A;Residues: 157-505 <HRA>  
A;Cross-references: EMBL:X59741  
R;Parranen, J.; Maekela, T.P.; Alitalo, R.; Lehtvaeslahti, H.; Alitalo, K.  
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990  
A;Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.  
A;Reference number: A38268; MUID:91062389; PMID:2247464  
A;Accession: C38268  
A;Status: nucleic acid sequence not shown; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 362-417 <PAR>  
C;Genetics:  
A;Gene: GDB:RCK  
A;Cross-references: GDB:119303; OMIM:142370  
A;Map position: 20q11-20q12  
A;Introns: 207/1; 258/1; 318/1; 343/3; 395/1; 439/1  
C;Function:  
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phos  
F;2-505/Product: protein-tyrosine kinase hck #status predicted <MAT>  
F;64-112/Domain: SH3 homology <SH3>  
F;123-220/Domain: SH2 homology <SH2>  
F;239-497/Domain: protein kinase homology <KIN>  
F;247-255/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;3/Binding site: palmitate (Cys) (covalent) #status predicted  
F;269/Active site: Lys #status predicted  
F;390/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 26.6%; Score 360.5; DB 1; Length 505;  
Best Local Similarity 41.6%; Pred. No. 4.6e-22;  
Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;

QY 12 PPSLSVSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPLTIVSEDGDMTV 71  
Db 40 PGPNSHNS- --NTPGLREAGSDDIIWALYDYEAITHHEDLSFKGDDQMVVLESGEWWKA 96  
QY 72 LSEVSGREYNIPSVHGVK- ---SHGWLYEGLSRKAEELLILLPGNPGAPLIRSQTER 127  
Db 97 RSLATRKEGYIPSNYVARVDSLETBEWFFKISRKDAERQLLAPGNMLGSPMIRDSSETTK 156  
QY 128 GSYLSVRLSRPASWDRIHRHYRIHCLDNGWLYISPRITFPPSLQALVDHYSLEADICCLL 187  
Db 157 GSYLSVSDYDPRQGTGVKHYKIRTLNDGGFFISPRITFSTLQELVDHYKNGDGLCCKL 216

RESULT 6  
JQJ321  
protein-tyrosine kinase (EC 2.7.1.112) hck - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 04-Feb-2000  
C;Accession: JQJ321; S18974  
R;Okano, Y.; Sugimoto, Y.; Fukuoka, M.; Matsui, A.; Nagata, K.; Nozawa, Y.  
Biochem. Biophys. Res. Commun. 181, 1137-1144, 1991  
A;Title: Identification of rat cDNA encoding hck tyrosine kinase from megakaryocytes.  
A;Reference number: JQJ321; MUID:92109719; PMID:1764064

A:Accession: J01321  
A:Molecule type: mRNA  
A:Residues: 1-503 <OKA>  
A:Cross-references: GB:S74141; NID:G241436; PIDN:AAB20754.1; PID:G241437  
A:Experimental source: megakaryocyte  
R:Rena, V.; Swarup, G.  
Submitted to the EMBL Data Library, December 1991  
A:Reference number: S18974  
A:Accession: S18974  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-50, 'V', 52-204, 'R', 206-305, 'T', 307-503 <REM>  
A:Cross-references: EMBL:X62345; NID:G57581; PIDN:CAA44218.1; PID:G57582  
C:Genetics:  
A:Gene: hck  
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology; SH3 h  
C:Keywords: ATP; autophosphorylation; blocked amino end; kinase-related transforming pro  
n kinase  
F:62-110/Domain: SH3 homology <SH3>  
F:121-218/Domain: SH2 homology <SH2>  
F:237-495/Domain: protein kinase homology <KIN>  
F:245-253/Region: protein kinase ATP-binding motif  
F:245-253/Region: protein kinase ATP-binding motif  
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F:3/Binding site: palmitate (Cys) (covalent) #status predicted  
F:267/Active site: Lys #status predicted  
F:388/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 26.1%; Score 353.5; DB 1; Length 503;  
Best Local Similarity 37.1%; Pred. No. 1.7e-21;  
Matches 85; Conservative 36; Mismatches 77; Indels 31; Gaps 5;  
QY 1 MGSLSR--RKSLPSPSSSVQGGPVMT-----AERSKAT-A 37  
DB 1 MGVKSRRLRGSKASKLEPNANOKGVYVDPDTPSKKLPNSINSPPGFVEGSEDTIV 60  
QY 38 VALGSPAGGPAELSLRIGELPLTIVSEDDWMTVLSEVSGREYNIPSVHVGK-----SHG 93  
DB 61 VALDYIEAIHREDISFQGGDQWVLSEGEWKAARSATKKEGIPSNYVARVNSLETET 120  
QY 94 WLYEGLSREKABELLLPGNPGGAPLIRSEOTRRGSYSLSVRLSRPASWDRIHYRIHCL 153  
DB 121 WFFKGISRKAERHLLAPGNMLGSMIRDSSTTKGYSLSVRDPPQHGDTVKHYKIRT 180  
QY 154 DNGWLYISPRITFSPSLQALVDHYSELADDICLLKEPCVLORAGPLPK 202  
DB 181 DSGGFIYISPRSTFSSQLVHVKKGDLCKLVPCV-----SPKPK 225

RESULT 7  
137206  
Protein-tyrosine kinase (EC 2.7.1.112) blk - human  
C:Species: Hmo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 04-Mar-2000  
A:Accession: I37206; S51647  
R:Islam, K.B.; Rabbani, H.; Larsson, C.; Sanders, R.; Smith, C.I.  
J. Immunol. 154, 1265-1272, 1995  
A:Title: Molecular cloning, characterization, and chromosomal localization of a human ly  
A:Reference number: 137206; MUID:95123078; PID:7822795  
A:Accession: I37206  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-505 <RES>  
A:Cross-references: EMBL:Z33998; NID:G601951; PIDN:CAA83965.1; PID:G601952  
C:Genetics:  
A:Gene: GDB:BLK  
A:Cross-references: GDB:454114; OMIM:191305  
A:Map position: 8p23-8p22  
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C:Keywords: ATP; blocked amino end; lipoprotein; myristylation; phosphotransferase; tyro  
F:65-113/Domain: SH3 homology <SH3>  
F:124-220/Domain: SH2 homology <SH2>  
F:239-497/Domain: protein kinase homology <KIN>  
F:247-255/Region: protein kinase ATP-binding motif

F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F:269/Active site: Lys #status predicted

Query Match 26.1%; Score 352.5; DB 2; Length 505;  
Best Local Similarity 43.6%; Pred. No. 2.1e-21;  
Matches 75; Conservative 24; Mismatches 68; Indels 5; Gaps 2;  
QY 25 PVTMEASRSKATAVALGSPAGGPAELSLRIGELPLTIVSEDDWMTVLSEVSGREYNIPS 84  
DB 51 PPDEHLDCKHFVALYDYTTAMNDRDLQMLKGEKLOVLKGTGDWMLARSLVTGEGYPS 110  
QY 85 VHVGVKVS----HGMWLEGLSREKABELLLPGNPGGAPLIRSEOTRRGSYSLSVRLSRPA 140  
DB 111 NFVAVESLESMEWPFRRQGRKEAERQLLAPINKAGSFLIRESETKNGAFSLVK-DVTT 169  
QY 141 SWDRIRHYRIHCLDNGWLYISPRITFSPSLQALVDHYSELADDICLLKEPCV 192  
DB 170 QGELIKHYKIRCLDEGGYIIPRIITFSPSLQALVDHYSKGDLQORLTPLCV 221

RESULT 8  
TVMSHC  
Protein-tyrosine kinase (EC 2.7.1.112) hck - mouse  
N:Alternate names: kinase-related transforming protein (bmk)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 28-Jan-2000  
A:Accession: A27282; A39973  
R:Klemenz, M.J.; McKeircher, S.R.; Maki, R.A.  
Nucleic Acids Res. 15, 9600, 1987  
A:Title: Nucleotide sequence of the mouse hck gene.  
A:Reference number: A27282; MUID:88067781; PMID:3684607  
A:Accession: A27282  
A:Molecule type: mRNA  
A:Residues: 1-503 <KLE>  
A:Cross-references: GB:Y00487; NID:G51209; PIDN:CAA68544.1; PID:G51210  
R:Holtzman, D.A.; Cook, W.D.; Dunn, A.R.  
Proc. Natl. Acad. Sci. U.S.A. 84, 8325-8329, 1987  
A:Title: Isolation and sequence of a cDNA corresponding to a src-related gene expressed i  
A:Reference number: A39973; MUID:88068587; PMID:3317404  
A:Molecule type: mRNA  
A:Status: preliminary; not compared with conceptual translation  
A:Residues: 1-503 <HOL>  
A:Cross-references: GB:J03023; NID:G192212; PIDN:AAA37305.1; PID:G309118  
C:Genetics:  
A:Gene: hck  
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phos  
F:62-110/Domain: SH3 homology <SH3>  
F:121-218/Domain: SH2 homology <KIN>  
F:237-495/Domain: protein kinase homology <KIN>  
F:245-253/Region: protein kinase ATP-binding motif  
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F:3/Binding site: palmitate (Cys) (covalent) #status predicted  
F:267/Active site: Lys #status predicted  
F:388,499/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

Query Match 26.0%; Score 352; DB 1; Length 503;  
Best Local Similarity 41.0%; Pred. No. 2.3e-21;  
Matches 82; Conservative 33; Mismatches 73; Indels 12; Gaps 4;  
QY 11 LPSLSLSSSVQGGPVMT-----AERSKAT-AVALGSPAGGPAELSLRIGELPLTIVSE 66  
DB 30 VPDPTSSSKLGNNSNMPFGFVEGSEDTIVVALYDYIEAIHREDLSFGQDQWVLEAG 89  
QY 67 DWMTVLSEVSGREYNIPSVHVGK-----SHGWLSEGLSREKABELLLPGNPGGAPLIR 122  
DB 90 EHWKARSATKKEGIPSNYVARVNSLETETWFFKISRKAERHLLAPGNMLGSMIR 149  
QY 123 SOTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFSPSLQALVDHYSELADD 182  
DB 150 SETTKGSYSLSVRDPPQHGDTVKHYKIRTLDSGGFIYISPRSTFSSQLVHVKKG 209

QY 183 ICCLKEPCVQLQAGPLGK 202  
Db 210 LCQKLSVPCV----SPKPK 225

RESULT 9  
A40092  
protein-tyrosine kinase (EC 2.7.1.112) blk [validated] - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 16-Jun-2000 #sequence\_revision 16-Jun-2000 #text\_change 16-Jun-2000  
C:Accession: A40092  
R:Dymekki, S.M.; Niederhuber, J.E.; Desiderio, S.V.  
Science 247, 332-336, 1990  
A>Title: Specific expression of a tyrosine kinase gene, blk, in B lymphoid cells.  
A:Reference number: A40092; MUID:90117147; PMID:2404338  
A:Accession: A40092  
A:Molecule type: mRNA  
A:Residues: 1-499 <DIM>  
A:Cross-references: GB:M30903; NID:G202076; PIDN:AAA40453.1; PID:G202077  
C:Genetics:  
A:Gene: MGI:Blk  
A:Map position: 14:28.0  
A:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology  
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho  
F:59-107/Domain: SH2 homology <SH2>  
F:118-214/Domain: SH2 homology <SH2>  
F:233-491/Domain: protein kinase homology <KIN>  
F:241-249/Region: protein kinase ATP-binding motif  
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F:263/Active site: Lys #status predicted

Query Match 25.5%; Score 344.5; DB 1; Length 499;  
Best Local Similarity 36.0%; Pred. No. 9.5e-21;  
Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;

QY 1 MGLSPRRKSLSPSLSSVQGGPV-----TWAEERSK 34  
Db 1 MGLSSKRQ-----VSEKGMKSPVKINTQKAPPLPLVFNHLAPPNPQDPDEE 54

QY 35 ATAVALGSPFAGGPAELSLRLGEPLTIVSEDGDWTVLSEVSGREYNIPSVHVGVKS-- 91  
Db 55 RFVVALFDYAANDRLQVLKQKQLVLRSTGDWMLARSLVTRGEGYVPSNFVAPVETLE 114

QY 92 -HGWLYEGLSRKAEELLNPNPGGAFIRSQTRRSYSLSVLSRSPASDRIRHVRI 150  
Db 115 VEKMFRTISRDAERQLLAPNWKAGSFLIRSESNNKGAFLSVK-DITTOGEVVVHYKI 173

QY 151 HCLDNGMLVSPRLTSPFQALVDHYSELADPICLLKEPCV 192  
Db 174 RSLDNGGYVISRIFPTTLQALVQHYSKGDLGCKLTLPCV 215

RESULT 10  
ORHULK  
N:Alternate names: kinase-related transforming protein (lck)  
C:Species: Homo sapiens (man)  
C>Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 28-Jan-2000  
C:Accession: JQ0152; S07822; S07200; S01879; S07143; A32797; I57636  
R:Rouer, E.; Van Huynh, T.; de Souza, S.L.; Lang, M.C.; Fischer, S.; Benarous, R.  
Gene 84, 105-113, 1989  
A>Title: Structure of the human lck gene: differences in genomic organisation within src  
A:Reference number: JQ0152; MUID:90108697; PMID:2558056  
A:Accession: JQ0152  
A:Molecule type: DNA  
A:Residues: 1-509 <ROU>  
A:Cross-references: EMBL:X14053  
R:Perlmutter, R.M.; Marth, J.D.; Lewis, D.B.; Peet, R.; Ziegler, S.F.; Wilson, C.B.  
J. Cell. Biochem. 38, 117-126, 1988  
A>Title: Structure and expression of lck transcripts in human lymphoid cells.  
A:Reference number: S07822; MUID:89123626; PMID:3265417  
A:Accession: S07822

A:Molecule type: mRNA  
A:Residues: 1-86, 'P', 88-509 <PER>  
A:Cross-references: EMBL:X13529; NID:G34294; PIDN:CAA31884.1; PID:G34295  
R:Koga, Y.; Caccia, N.; Toyonaga, B.; Spolski, R.; Yanagi, Y.; Yoshikai, Y.; Mak, T.W.  
Eur. J. Immunol. 16, 1643-1646, 1986  
A>Title: A human T cell-specific cDNA clone (YT16) encodes a protein with extensive homol  
A:Reference number: S07200; MUID:87133831; PMID:3493153  
A:Accession: S07200  
A:Molecule type: mRNA  
A:Residues: 1-205, 'ASAITPI', 212-257, 'RCGW', 262, 'TTT', 266, 'T', 268-281, 'AGRIPL', 287-503, 'ST'  
A:Cross-references: EMBL:X05027; NID:G36807; PIDN:CAA28691.1; PID:G36808  
R:Veillette, A.; Foss, F.M.; Sausville, E.A.; Bolen, J.B.; Rosen, N.  
Oncogene Res. 1, 357-374, 1987  
A>Title: Expression of the lck tyrosine kinase gene in human colon carcinoma and other nc  
A:Reference number: S01879; MUID:88217332; PMID:2835736  
A:Accession: S01879  
A:Molecule type: mRNA  
A:Residues: 368-471, 'H', 473-509 <VEI>  
A:Cross-references: EMBL:X06369; NID:G34288; PIDN:CAA29667.1; PID:G34289  
R:Trevillian, J.M.; Lin, Y.; Chen, S.J.; Phillips, C.A.; Canna, C.; Linna, T.J.  
Biochim. Biophys. Acta 888, 286-295, 1986  
A>Title: Human T lymphocytes express a protein-tyrosine kinase homologous to p56 (LSTRA).  
A:Reference number: S07143; MUID:87000726; PMID:3489436  
A:Accession: S07143  
A:Molecule type: mRNA  
A:Residues: 'A', 376-509 <TRE>  
A:Cross-references: EMBL:X04476; NID:G35779; PIDN:CAA28165.1; PID:G35780  
R:Takadera, T.; Leung, S.; Gernone, A.; Koga, Y.; Takiyama, Y.; Miyamoto, N.G.; Mak, T.W.  
Mol. Cell. Biol. 9, 2173-2180, 1989  
A>Title: Structure of the two promoters of the human lck gene: differential accumulation  
A:Reference number: A32797; MUID:89313764; PMID:2787474  
A:Accession: A32797  
A:Molecule type: DNA  
A:Residues: 1-35 <TAX>  
A:Cross-references: GB:M26692; NID:G341523; PIDN:AAA59503.1; PID:G349702  
R:Garvin, A.M.; Pawar, S.; Marth, J.D.; Perlmutter, R.M.  
Mol. Cell. Biol. 8, 3058-3064, 1988  
A>Title: Structure of the murine lck gene and its rearrangement in a murine lymphoma cell  
A:Reference number: I57636; MUID:89096891; PMID:2850479  
A:Accession: I57636  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-35, 'VR', <RES>  
A:Cross-references: GB:M21510; NID:G187031; PIDN:AAA59501.1; PID:G553522  
C:Comment: Protein tyrosine kinases play important roles in the control of cell growth ar  
C:Genetics:  
A:Gene: GDB:LCK  
A:Cross-references: GDB:119360; OMIM:153390  
A:Map position: 1p35-1p34.3  
A:Introns: 35/3; 63/1; 93/2; 126/2; 161/1; 211/1; 262/1; 322/1; 347/3; 399/1; 443/1  
C:Function:  
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 ho  
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho  
F:2-509/Product: protein-tyrosine kinase lck #status predicted <MAT>  
F:68-116/Domain: SH3 homology <SH3>  
F:127-224/Domain: SH2 homology <SH2>  
F:243-501/Domain: protein kinase homology <KIN>  
F:251-259/Region: protein kinase ATP-binding motif  
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F:3.5/Binding site: palmitate (Cys) (covalent) #status predicted  
F:273/Active site: Lys #status predicted  
F:394,505/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

Query Match 25.1%; Score 340; DB 1; Length 509;  
Best Local Similarity 40.6%; Pred. No. 2.3e-20;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

QY 26 VTWEAERSKAT-----AVALGSPFAGGPAELSLRLGEPLTIVSEDGDWTVLSEVSGRE 79  
Db 49 VTVEGSPFPASPQDNLVIALHSYFSDGDLGFEQGLRILEQGEWKKASLTIGQE 108

QY 80 YNIPSVHVGVKS-----HGWLYEGLSRKAEELLNPNPGGAFIRSQTRRSYSLSVR 135

Db 109 GIPFNFAKNSLEPEPWFKNLSRKAERQQLAPGNTHGSFLRSESTAGSFLSVR 168  
QY 136 LSRPASWRIRHYRICHLDNGWLYSPRLTPPSLQALVDHYSELADIDICLLKPEPCVLQR 195  
Db 169 DFDQNGQEVVHKYKRNLDNGGFIYSPRITPFGHLYVHTNASDGLCTSLSPCQTQK 228

RESULT 11  
I48845  
protein-tyrosine kinase (EC 2.7.1.112) lck, lymphocyte - mouse  
N/Alternate names: p56; protein-tyrosine kinase tck  
C/Species: Mus musculus (house mouse)  
C/Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 03-Mar-2000  
C/Accession: I48845; A23639; I57629; I77452  
R/Voronova, A.F.; Sefton, B.M.  
Nature 319, 682-685, 1986  
A/Title: Expression of a new tyrosine protein kinase is stimulated by retrovirus promote  
A/Reference number: I48845, MUID:86146842; PMID:3081813  
A/Accession: I48845  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-509 <VORI>  
A/Cross-references: EMBL:X03533; NID:G54813; PIDN:CAA27234.1; PID:G54814  
R/Marsh, J.D.; Peet, R.; Krebs, E.G.; Perlmutter, R.M.  
Cell 43, 393-404, 1985  
A/Title: A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpres  
A/Reference number: A23639; MUID:86079521; PMID:2416464  
A/Accession: A23639  
A/Molecule type: mRNA  
A/Residues: 1-282, 'VP', 285-509 <VAR>  
A/Cross-references: GB:M12056; NID:G198763  
A/Note: the sequence is revised in GenBank entry MUSLCK, release 116.0, (PIDN:AAB59674.1  
R/Voronova, A.F.; Adler, H.T.; Sefton, B.M.  
Mol. Cell. Biol. 7, 4407-4413, 1987  
A/Title: Two lck transcripts containing different 5' untranslated regions are present in  
A/Reference number: I57629; MUID:86142832; PMID:3501824  
A/Accession: I57629  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-11 <VOR>  
A/Cross-references: GB:M18098; NID:G198766; PIDN:AAA39421.1; PID:G198767  
R/Garvin, A.M.; Pawar, S.; Marsh, J.D.; Perlmutter, R.M.  
Mol. Cell. Biol. 8, 3058-3064, 1988  
A/Title: Structure of the murine lck gene and its rearrangement in a murine lymphoma cel  
A/Reference number: I57636; MUID:89096891; PMID:2850479  
A/Accession: I77452  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-35, 'VR', <GAR>  
A/Cross-references: GB:M21511; NID:G198768; PIDN:AAA39422.1; PID:G554186  
C/Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C/Keywords: ATP; autophosphorylation; blocked amino end; kinase-related transforming pro  
F/68-116/Domain: SH3 homology <SH3>  
F/127-224/Domain: SH2 homology <SH2>  
F/243-501/Domain: protein kinase homology <KIN>  
F/251-259/Region: protein kinase ATP-binding motif  
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F/273/Active site: Lys #status predicted  
F/394,505/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 25.0%; Score 338.5; DB 1; Length 509;  
Best Local Similarity 39.1%; Pred. No. 3e-20;  
Matches 77; Conservative 28; Mismatches 85; Indels 7; Gaps 2;

QY 3 SLPSRRKSLPSSLSVSGQGVPTWEAERSKATAVALGSPFAGGPAELSLRLGPEPLTIV 62  
Db 35 SLPIRNGSEVRDPL---VTYEGSLPPASPLQNLVIALHSYRPFSDGLGFEGKQGLRIL 91  
QY 63 SEDGDWMTVLSEVSGREYNIPSVHVGKVS---HGWLYEGLSRKAEELLILPGNPGGAF 118  
Db 92 EQSGEWWKQSLTTGGEGIPFNFAKNSLEPEPWFKNLSRKAERQQLAPGNTHGSF 151

QY 119 LIRESTREGSYLSVRLSRPASWDRIRHYRICHLDNGWLYSPRLTPPSLQALVDHYSE 178  
Db 152 LIRESESTAGSFLSVRLSRPASWDRIRHYRICHLDNGWLYSPRLTPPSLQALVDHYSE 211  
QY 179 LADDICLLKPEPCVLQR 195  
Db 212 ASDGLCTKLSRCQTQK 228

RESULT 12  
A39939  
protein-tyrosine kinase (EC 2.7.1.112) tk1 [similarity] - chicken  
N/Alternate names: kinase-related transforming protein (tk1); T-cell surface antigen ass  
C/Species: Gallus gallus (chicken)  
C/Date: 16-Jun-2000 #sequence\_revision 16-Jun-2000 #text\_change 16-Jun-2000  
C/Accession: A42126; A39939  
R/Chow, L.M.; Ratcliffe, M.J.; Veillette, A.  
Mol. Cell. Biol. 12, 1226-1233, 1992  
A/Title: tk1 is the avian homolog of the mammalian lck tyrosine protein kinase gene.  
A/Reference number: A42126; MUID:92186854; PMID:1545804  
A/Accession: A42126  
A/Molecule type: mRNA  
A/Residues: 1-88 <CHC>  
A/Cross-references: GB:M85043  
A/Experimental source: thymus, spleen  
A/Note: sequence extracted from NCBI backbone (NCBIN:88831, NCBI:88833)  
R/Srebbardt, K.; Mullins, J.I.; Bruck, C.; Ruebaamen-Waigmann, H.  
Proc. Natl. Acad. Sci. U.S.A. 84, 8778-8782, 1987  
A/Title: Additional member of the protein-tyrosine kinase family: the src-and lck-related  
A/Reference number: A39939; MUID:88097370; PMID:3321053  
A/Accession: A39939  
A/Molecule type: mRNA  
A/Residues: 52-507 <STR>  
A/Cross-references: GB:J03579; NID:G212712; PIDN:AAA49081.1; PID:G212713  
C/Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C/Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho  
F/65-114/Domain: SH3 homology <SH3>  
F/125-222/Domain: SH2 homology <SH2>  
F/241-499/Domain: protein kinase homology <KIN>  
F/249-257/Region: protein kinase ATP-binding motif  
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F/392,503/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 24.6%; Score 333; DB 1; Length 507;  
Best Local Similarity 43.3%; Pred. No. 8.7e-20;  
Matches 70; Conservative 27; Mismatches 61; Indels 4; Gaps 1;

QY 38 VALGSPFAGGPAELSLRLGCEPLTIVSGDGDWMTVLSEVSGREYNIPSVHVGKVS----HG 93  
Db 65 VALYDYEPTHDGLGLKQGEKLRVLSESGEWRAQSLTTGQEGLIPIHNFVAMVNSLEPEP 124  
QY 94 WLVEGLSRKAEELLILPGNPGGAFILRESQTRRGSYLSVRLSRPASWDRIRHYRICH 153  
Db 125 WFFKNLSRKNAAERLASNGTHGSFLIRSETSKGSYLSVRLSRPASWDRIRHYRICH 184  
QY 154 DNGWLYSPRLTPPSLQALVDHYSELADIDICLLKPEPCVLQR 195  
Db 185 DNGYVLSRVTSSHLHELVEYSSSDGLCTSLGKPCQTQK 226

RESULT 13  
B49114  
protein-tyrosine kinase (EC 2.7.1.112) fvk - Pacific electric ray  
C/Species: Torpedo californica (Pacific electric ray)  
C/Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 18-Feb-2000  
C/Accession: B49114  
R/Swope, S.L.; Huganir, R.L.  
J. Biol. Chem. 268, 25152-25161, 1993  
A/Title: Molecular cloning of two abundant protein tyrosine kinases in Torpedo electric  
A/Reference number: A49114; MUID:94043386; PMID:8227079  
A/Accession: B49114  
A/Status: preliminary  
A/Molecule type: mRNA



A;Residues: 1-539 <SMO>

A;Cross-references: GB:U01350  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho  
F;91-140/Domain: SH3 homology <SH3>  
F;151-248/Domain: SH2 homology <SH2>  
F;271-529/Domain: protein kinase homology <KIN>  
F;279-287/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;301/Active site: Lys #status predicted  
F;422,533/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 24.0%; Score 325; DB 2; Length 539;  
Best Local Similarity 41.0%; Pred. No. 4.3e-19;  
Matches 75; Conservative 27; Mismatches 67; Indels 14; Gaps 4;  
QY 15 SLSSVVG-OGPTVMEARSKATAVALGSPFAGPAELSLRLGEPITV-SEDGDMWTVL 72  
Db 74 SLTSRVGGVGGVTM-----FIALYNYDARTEDLTFKRGKFXHINSSEGDWEAR 125  
QY 73 SEVSGREYNIPSVHVGK-----SHGWLVEGLGREKAEELLLPGNPGAFIIRSQTRRG 128  
Db 126 SLITGSGTGYIPSVNAPVDSIQAEWYFGKMGKOTERMLLCPGNPRGTFLIRESETKG 185  
QY 129 SYSLSVLPSPASWDRIHRVHCLDNGWLYISPLTPPSQLQVLDHSEADICLLK 189  
Db 186 AYSLSIRDWDEKGDHVGKIRKLDNGGYITRAQFETVQVHHVTEBAAGLCRLV 245  
QY 189 EPC 191  
Db 246 VPC 248

## RESULT 14

TVFVG9

Protein-tyrosine kinase (EC 2.7.1.112) yes - avian sarcoma virus Y73

C;Species: avian sarcoma virus Y73  
A;Note: host Gallus gallus (chicken)  
C;Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 23-Feb-1997  
C;Accession: A00633  
R;Kitamura, N.; Kitamura, A.; Toyoshima, K.; Hirayama, Y.; Yoshida, M.  
Nature 297, 205-208, 1982  
A;Title: Avian sarcoma virus Y73 genome sequence and structural similarity of its transp  
A;Reference number: A00633; MUID:82195528; PMID:6281656  
A;Molecule type: genomic RNA  
A;Residues: 1-528 <KIT>  
C;Comment: This protein is synthesized as a gag-yes polypeptide.  
C;Genetics:  
A;Gene: yes  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; tran  
F;88-137/Domain: SH3 homology <SH3>  
F;148-245/Domain: SH2 homology <SH2>  
F;265-523/Domain: protein kinase homology <KIN>  
F;273-281/Region: protein kinase ATP-binding motif  
F;295/Active site: Lys #status predicted  
F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 23.7%; Score 320; DB 1; Length 528;  
Best Local Similarity 31.5%; Pred. No. 1.1e-18;  
Matches 91; Conservative 43; Mismatches 105; Indels 50; Gaps 9;  
QY 2 GSLPSRRKSLPSPSSVGGQGVPTMEARSKATAVALGSPFAGPAELSLRLGEPITV 61  
Db 62 GGASSSPSPSP-YPSTLTGGTV-----FVALDYEARITDLSFKKGERFQI 110  
QY 62 VSE-DGDWTVLSEVSGREYNIPSVHVGK-----SHGWLVEGLGREKAEELLLPGNPG 116  
Db 111 INNTEGDWNEARSIATGKTYIFSNVTPADSIQAEWYFGKMGKOTERMLLCPGNPRGTFLIRESETKG 170  
QY 117 AFLIREQTRGYSVLSVRSPASWDRI-----HYRHLCDNGWLYISPLTPPSQLQ 171

Db 171 IFLVRESETTKGYSLISIR-----DWDEVRGDNVVKYKIRKLDNGGYITTRAQFESLQK 225  
QY 172 LVDPVHSEADICLLKEPC-----VLQAGPLPGKDIPLVTVQR----- 212  
Db 226 LVKHREHADGLCHKLTVCTVTPQOGLAKDAWEIPRESLRLEVKLGCGCFGEVWMT 285  
QY 213 -----TFLNWKELDSLLFSEAAATGEESLLSEGURESL-SFYISLNDVAV 256  
Db 286 WNGTTTKVAIKTLKLTGTMPEAFLOEAQIMKLRHDKLPLVLYAVVSEPI 334

## RESULT 15

A45501  
Protein-tyrosine kinase (EC 2.7.1.112) yes [similarity] - African clawed frog  
N;Alternate names: kinase-related transforming protein (yes)  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 16-Jun-2000 #sequence\_revision 16-Jun-2000 #text\_change 16-Jun-2000  
C;Accession: A45501; S08517  
R;Steele, R.E.; Irwin, M.Y.; Knudsen, C.L.; Collett, J.W.; Fero, J.B.  
Oncogene Res. 1, 223-233, 1989  
A;Title: The yes proto-oncogene is present in amphibians and contributes to the maternal  
A;Reference number: A45501  
A;Accession: A45501  
A;Molecule type: mRNA  
A;Residues: 1-537 <STE>  
A;Cross-references: GB:X14377  
R;Steele, R.E.; Irwin, M.Y.; Knudsen, C.L.; Collett, J.W.; Fero, J.B.  
submitted to the EMBL Data Library, February 1989  
A;Reference number: S08517  
A;Accession: S08517  
A;Molecule type: mRNA  
A;Residues: 1-250, 'S', 252-537 <ST2>  
A;Cross-references: EMBL:X14377; MID:g65272; PIDN:CAA32551.1; FID:g65273  
C;Genetics:  
A;Gene: yes  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: ATP; autophosphorylation; blocked amino end; kinase-related transforming prot  
F;92-141/Domain: SH3 homology <SH3>  
F;152-249/Domain: SH2 homology <SH2>  
F;269-527/Domain: protein kinase homology <KIN>  
F;277-285/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;299/Active site: Lys #status predicted  
F;420,531/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

Query Match 23.6%; Score 319.5; DB 1; Length 537;  
Best Local Similarity 33.2%; Pred. No. 1.2e-18;  
Matches 84; Conservative 38; Mismatches 92; Indels 39; Gaps 7;

QY 38 VALGSPFAGGPAELSLRLGEPITVSE-DGDWTVLSEVSGREYNIPSVHVGK-----SH 92  
Db 91 VALDYEARITDLSFKKGERFQIINTEGDWNEARSIATGKTYIFSNVTPADSIQAE 150  
QY 93 GWLYEGLGREKAEELLLPGNPGAFIIRSQTRGYSLSVLSRSPASWDRI-----H 147  
Db 151 EWFYFGKMGKDAERLLNPGNQGTFLVRESETTKGYSLISIR-----DWDEVRGDNVVKH 205  
QY 148 YRHLCDNGWLYISPLTPPSQLQVLDHSEADICLLKEPC-----VLQAGP 198  
Db 206 YKIRKLDNGGYITTRAQFESLQKLVKHSEHADGLCYRLTTCVTKVQTOGLAKDAWE 265  
QY 199 LPKDIPLVTVQR-----TFLNWKELDSLLFSEAAATGEESLLSEGURESL 244  
Db 266 IPRESLRDLVKLGCGCFGEVWIGTWNTGTTKVAIKTLKPGTMPEAFLOEAQIMKLRHDK 325  
QY 245 L-SFYISLNDVAV 256  
Db 326 LVPLVAVVSEPI 338

Search completed: February 20, 2004, 09:33:04  
Job time : 46 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 08:16:06, Search time 39 Seconds  
(without alignments)  
348.469 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLPSRKPSLSSV.....RESLSFYISUNDEAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1342	99.2	260	1 SLA2_HUMAN	Qh6q3 homo sapien
2	1023	75.6	258	1 SLA2_MOUSE	Q8r410 mus musculus
3	501	37.0	312	1 SLA1_RAT	P59622 rattus norv
4	491.5	36.3	280	1 SLA1_MOUSE	Q60898 mus musculus
5	481.5	35.6	275	1 SLA1_HUMAN	Q13239 homo sapien
6	370.5	27.4	511	1 LYN_HUMAN	P07948 homo sapien
7	367.5	27.2	511	1 LYN_MOUSE	P25911 mus musculus
8	367.5	27.2	511	1 LYN_RAT	Q27014 rattus norv
9	360.5	26.6	525	1 HCK_HUMAN	P08631 homo sapien
10	358	26.5	503	1 HCK_MACFA	Q8sm30 macaca fasc
11	352.5	26.1	504	1 BLK_HUMAN	Q51451 homo sapien
12	352	26.0	523	1 HCK_MOUSE	P08103 mus musculus
13	351.5	26.0	502	1 HCK_RAT	P50545 rattus norv
14	341.5	25.2	498	1 BLK_MOUSE	P16277 mus musculus
15	340	25.1	508	1 LCK_HUMAN	P06239 homo sapien
16	338.5	25.0	508	1 LCK_MOUSE	P06240 mus musculus
17	333	24.6	507	1 LCK_CHICK	P22683 gallus gall
18	332	24.5	517	1 SR42_DROME	Q9v9j3 grosophila
19	319.5	23.6	537	1 YES_XENLA	P10936 xenopus lae
20	317	23.4	541	1 YES_CHICK	P09324 gallus gall
21	315.5	23.3	536	1 FYN_HUMAN	P06241 homo sapien
22	315.5	23.3	543	1 YES_HUMAN	P07947 homo sapien
23	312.5	23.1	529	1 FGR_HUMAN	P09769 homo sapien
24	311	23.0	544	1 YES_XPHE	P27447 xiphophorus
25	310.5	22.9	536	1 FYN_XENLA	P13405 xenopus lae
26	309	22.8	528	1 YES_AVISR	P00527 avian sarco
27	309	22.8	533	1 FYN_MOUSE	P39689 mus musculus
28	306.5	22.7	541	1 YES_MOUSE	Q04736 mus musculus
29	305.5	22.6	536	1 FYN_XPHE	P27446 xiphophorus
30	303	22.4	517	1 FGR_MOUSE	P14234 mus musculus
31	301	22.2	535	1 YRK_CHICK	Q02977 gallus gall
32	300	22.2	539	1 YES_CANFA	Q28923 canis famil
33	297.5	22.0	533	1 FYN_CHICK	Q05876 gallus gall

34	290	21.4	506	1 SRK4_SPOLA	P42890 sporgilla 1
35	286.5	21.2	526	1 SRC_AVISR	P00525 avian sarco
36	286.5	21.2	532	1 SRC_CHICK	P00523 gallus gall
37	286.5	21.2	552	1 SR64_DROME	P00528 drosophila
38	286.5	21.2	557	1 SRC_AVISR	P14085 avian sarco
39	286.5	21.2	587	1 SRC_AVIS2	P15054 avian sarco
40	284.5	21.0	568	1 SRC_AVIS3	P14084 avian sarco
41	277.5	20.5	535	1 SRC_RAT	Q9vud9 rattus norv
42	276.5	20.4	526	1 SRC_RSVSR	P00324 rous sarcom
43	276.5	20.4	535	1 SRC_HUMAN	P12931 homo sapien
44	276	20.4	509	1 STK_HYDRA	P17713 hydra atten
45	274.5	20.3	531	1 SRC2_XENLA	P13116 xenopus lae

## ALIGNMENTS

RESULT 1  
SLA2\_HUMAN  
ID SLA2\_HUMAN STANDARD; PRT; 260 AA.  
AC Q9H6Q3; Q8WY18; Q96Q14; Q9H135;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Src-like-adaptor 2 (Src-like adaptor protein-2) (SLAP-2) (Modulator of  
DE antigen receptor signaling) (MARS).  
GN SLA2 OR SLAP2 OR C20CRF156.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), CHARACTERIZATION, FUNCTION  
RP MYRISTYLATION, INTERACTION WITH CBL, AND MUTAGENESIS OF GLY-1.  
RX MEDLINE=21553259; PubMed=11696592;  
RA Holland S.J., Liao X.C., Mendenhall M.K., Zhou X., Pardo J., Chu P.,  
RA Spencer C., Fu A.C., Sheng N., Yu P., Pali E., Nagin A., Shen M.,  
RA Yu S., Chan E., Wu X., Li C., Wolsetschlager M., Aversa G.,  
RA Kolbinger F., Bennett M.K., Molineaux S., Luo Y., Payan D.G.,  
RA Mancebo H.S., Wu J.,  
RT "functional cloning of Src-like adapter protein-2 (SLAP-2), a novel  
RT inhibitor of antigen receptor signaling."  
J. Exp. Med. 194:1263-1276 (2001).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND ALTERNATIVE  
RP INITIATION.  
RC TISSUE=Thymus;  
RX MEDLINE=22415750; PubMed=12527895;  
RA Lorsto M.P., McGlade C.J.,  
RT "Cloning and characterization of human Src-like adaptor protein 2 and  
RT a novel splice isoform, SLAP-2-v."  
Oncogene 22:266-273 (2003).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Hepatoma;  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Ora T., Suzuki Y., Ohayashi M., Nishi T., Shibahara T.,  
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
RT "NEDD human cDNA sequencing project."  
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blake S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Comor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Hazley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehtvasilho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McComachie L.J., McDay K., McMurray A.,  
RA Milne S.A., Mastry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
RA Skuce C.D., Smith M.I., Soderlund C., Steward C.A., Sultun J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Walli M., Wallis J.M., Williams S.A.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20";  
RL Nature 414:865-871(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate; PubMed=12477932;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Klausner R.D., Collins F.S., Wagner L.H., Grouse L.H., Derge J.G.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RX CHARACTERIZATION.  
RP MEDLINE=22013997; PubMed=11891219;  
RA Pandey A., Ibarrola N., Kratchmarova I., Fernandez M.M.,  
RA Constantinescu S.N., Ohara O., Sawasdikosol S., Lodish H.F., Mann M.,  
RT "A novel Src homology 2 domain-containing molecule, Src-like adapter  
RT protein-2 (SLAP-2), which negatively regulates T cell receptor  
RT signaling";  
RL J. Biol. Chem. 277:19131-19138(2002).  
CC -!- FUNCTION: Adapter protein, which negatively regulates T-cell  
CC receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced  
CC activation of nuclear factor of activated T-cells. May act by  
CC linking signaling proteins such as ZAP70 with CBL, leading to a  
CC CBL dependent degradation of signaling proteins.  
CC -!- SUBUNIT: Interacts with phosphorylated proteins ZAP70 and CD32 via  
CC its SH2 domain (By similarity). Interacts with phosphorylated CBL  
CC via its C-terminal domain.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Isoform 1 is localized to the  
CC plasma membrane and intracellular vesicles, including late  
CC endosomal vesicles. Isoform 2 may be cytoplasmic and is not  
CC localized to membranes.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1; Synonyms=sp28;  
CC IsoId=Q9H6Q3-1; Sequence=Displayed;  
CC Note=Isoform 3 is produced by alternative initiation at  
CC Met-27 of isoform 1;  
CC Name=2; Synonyms=sp23, SLAP-2-V, MAR8-V;  
CC IsoId=Q9H6Q3-2; Sequence=VSP\_007240, VSP\_007241;  
CC Note=Isoform 4 is produced by alternative initiation at  
CC Met-27 of isoform 2;  
CC Event=Alternative initiation;

CC Comment=4 isoforms, 1/p28 (shown here), 2/p23, 3/p25 and 4/p20,  
CC ARE produced by alternative initiation at Met-0 and Met-27;  
CC -!- TISSUE SPECIFICITY: Predominantly expressed in immune system, with  
CC highest levels in peripheral blood leukocytes. Expressed in  
CC spleen, thymus and lymph nodes. Expressed in T-cells as well as in  
CC monocytes, and at low level in B-cells. Also detected in placenta,  
CC prostate, skin, retina and colon.  
CC -!- DOMAIN: The loss of the C-terminal domain partially abolishes the  
CC inhibitory function, but can be partially compensated by higher  
CC level of protein expression.  
CC -!- PTM: May be phosphorylated (By similarity).  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
CC -----  
CC EMBL; AF326353; AAL29204.1; -;  
CC EMBL; AF290985; AAL38197.1; -;  
CC EMBL; AF290986; AAL38198.1; -;  
CC EMBL; AK025645; BAB15201.1; -;  
CC EMBL; AL031662; CAC44645.1; ALT\_TERM.  
CC EMBL; AL050318; CAB75365.1; -;  
CC EMBL; BC042041; AAH42041.1; -;  
CC HSSP; P06239; ILKK.  
CC Genew; HGNC:17329; SLA2.  
CC MIM; 606577; -;  
CC InterPro; IPR000980; SH2.  
CC InterPro; IPR01452; SH3.  
CC Pfam; PF00017; SH2; 1.  
CC Pfam; PF00018; SH3; 1.  
CC PRINTS; PR0401; SH2DOMAIN.  
CC PRODOM; PD00093; SH2; 1.  
CC SMART; SM00252; SH2; 1.  
CC SMART; SM00326; SH3; 1.  
CC PROSITE; PS50001; SH2; 1.  
CC PROSITE; PS50002; SH3; 1.  
CC Membrane; SH2 domain; SH3 domain; Myristate; Phosphorylation;  
CC Alternative initiation; Alternative splicing; Lipoprotein.  
CC BY SIMILARITY.  
CC INIT MET 0 0  
CC CHAIN 1 260  
CC SRC-LIKE-ADAPTER 2, ISOFORM 1 AND ISOFORM  
CC 2.  
CC SRC-LIKE-ADAPTER 2, ISOFORM 3 AND ISOFORM  
CC 4.  
CC FOR ISOFORM 3 AND ISOFORM 4.  
CC SH3.  
CC SH2.  
CC SLA C-TERMINAL.  
CC N-myristoyl glycine (By similarity).  
CC LADDICLLKEPCVQLQAGPLGKQIPLPTV -> GWPAP  
CC WQGYPTTCDCAEDTTQLERAGLPVF (in isoform  
CC 2).  
CC /FTid-VSP\_007240.  
CC Missing (in isoform 2).  
CC /FTid-VSP\_007241.  
CC G->A: ABOLISHES LOCALIZATION TO  
CC MEMBRANES.  
CC SEQUENCE 260 AA; 28454 MW; A402C03449261B3D CRC64;  
CC -----  
CC Query Match 99.2%; Score 1342; DB 1; Length 260;  
CC Best Local Similarity 99.6%; Pred. No. 3.7e-105;  
CC Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
CC -----  
CC QY 2 GSLPSRKSLPSLSLSSVQGVPTWEARSKATAVALGSPAGGPAELSLRLSEPLTI 61  
CC DB 1 GSLPSRKSLPSLSLSSVQGVPTWEARSKATAVALGSPAGGPAELSLRLSEPLTI 60  
CC QY 62 VSEDGDWMTVLSEVSGREYNIPSVHGVKSHGWLVEGLSREKAEILLPLPGNGGAFATIR 121

Db 61 VSEDGDMWTVLSEVSREYNISVHVAKVSHGWLYEGLSRKAEELLLPGNPGAFLLR 120  
 QY 122 ESOTRGSGYSLVRLSRPASWDRIRHYRHCLDNGWLYISPLTPPSQALVDHYSELAD 181  
 Db 121 ESOTRGSGYSLVRLSRPASWDRIRHYRHCLDNGWLYISPLTPPSQALVDHYSELAD 180  
 QY 182 DICCLLKEPCVLQACGLPGKPIPLVTVQRTPLNWKELDSLLSEATGESLLSEGL 241  
 Db 181 DICCLLKEPCVLQACGLPGKPIPLVTVQRTPLNWKELDSLLSEATGESLLSEGL 240  
 QY 242 RESLSFYISLNDEAVSLDDA 261  
 Db 241 RESLSFYISLNDEAVSLDDA 260

RESULT 2  
 SLA2\_MOUSE STANDARD; PRT; 258 AA.  
 AC Q8R410; Q8COK2; Q8V142; Q9D1Z9;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Src-like-adaptor 2 (Src-like adaptor protein-2) (SLAP-2).  
 GN SLA2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), ALTERNATIVE INITIATION,  
 RP CHARACTERIZATION, FUNCTION, MYRISTOYLATION, PHOSPHORYLATION,  
 RP INTERACTION WITH ZAP70 AND CBL, AND MUTAGENESIS OF GLY-1; MET-26 AND  
 RP ARG-119.  
 RX MEDLINE=22022020; PubMed=12024036;  
 RA Loreto M.P., Berry D.M., McGlade C.J.;  
 RT "Functional cooperation between c-Chl and Src-like adaptor protein 2  
 RT in the negative regulation of T-cell receptor signaling.";  
 RL Mol. Cell. Biol. 22:4241-4255 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, MYRISTOYLATION,  
 RP INTERACTION WITH CBL; ZAP70 AND CD3Z, AND MUTAGENESIS OF GLY-1; PRO-81  
 RP ARG-119.  
 RX MEDLINE=22013997; PubMed=11891219;  
 RA Pandey A., Ibarrola N., Kratchmarova I., Fernandez M.M.,  
 RA Constantinescu S.N., Chira C., Sawaadikosol S., Lodish H.F., Mann M.;  
 RT "A novel Src homology 2 domain-containing molecule, Src-like adaptor  
 RT protein-2 (SLAP-2), which negatively regulates T cell receptor  
 RT signaling.";  
 RL J. Biol. Chem. 277:19131-19138 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=C57BL/6J, and NOD; TISSUE=Retina, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grammond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltas L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 CC -!- FUNCTION: Adapter protein, which negatively regulates T-cell  
 CC receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced  
 CC activation of nuclear factor of activated T-cells. May act by  
 CC linking signaling proteins such as ZAP70 with CBL, leading to a  
 CC CBL dependent degradation of signaling proteins.  
 CC -!- SUBUNIT: Interacts with phosphorylated CBL via its C-terminal  
 CC domain. Interacts with phosphorylated proteins ZAP70 and CD3Z via  
 CC its SH2 domain.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; localized to the plasma  
 CC membrane and intracellular vesicles, including late endosomal  
 CC vesicles.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Comment=2 isoforms, 1/p28 (shown here) and 2/p25, are produced  
 CC by alternative initiation at Met-0 and Met-26;  
 CC -!- TISSUE SPECIFICITY: Mainly expressed in immune system. Highly  
 CC expressed in spleen and thymus and expressed at intermediate  
 CC levels in lung. Not expressed in liver, heart and brain. Isoform 1  
 CC is predominant in lung and spleen, while isoform 2 is predominant  
 CC in thymus.  
 CC -!- DOMAIN: The loss of the C-terminal domain partially abolishes the  
 CC inhibitory function.  
 CC -!- PTM: May be phosphorylated.  
 CC -!- MISCELLANEOUS: Ref.1 confirmed the alternative initiation by  
 CC mutating the Met in position -1 to Val, and showed that isoform 1  
 CC is abolished in favor of isoform 2.  
 CC -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF287467; AAL38196.1; -;  
 CC EMBL; AF434990; AAL86403.1; -;  
 CC EMBL; AK020837; BAB32223.1; ALT\_INIT.  
 CC EMBL; AK030877; BAC37168.1; -;  
 CC EMBL; AK088672; BAC40495.1; -;  
 CC HSSP; P06239; 1LKK.  
 CC MGD; MGI:1925049; Sla2.  
 CC GO; GO:0016023; Cytoplasmic vesicle; IDA.  
 CC GO; GO:0005770; C:late endosome; IDA.  
 CC GO; GO:0005886; C:plasma membrane; IDA.  
 CC GO; GO:0005515; F:protein binding; IDA.  
 CC GO; GO:0042110; F:T-cell activation; IDA.  
 CC InterPro; IPR000980; SH2.  
 CC InterPro; IPR001452; SH3.  
 CC Pfam; PF00017; SH2; 1.  
 CC Pfam; PF00018; SH3; 1.  
 CC PRINTS; PR00401; SH2DOMAIN.  
 CC ProDom; PD000093; SH2; 1.  
 CC SMART; SM00252; SH2; 1.  
 CC SMART; SM00326; SH3; 1.  
 CC PROSITE; PS50001; SH2; 1.  
 CC PROSITE; PS50002; SH3; 1.  
 CC Immune response; Membrane; SH2 domain; SH3 domain; Myristate;  
 KW Phosphorylation; Alternative initiation; Lipoprotein.  
 FT INIT MET 0  
 FT CHAIN 1 258 SRC-LIKE-ADAPTER 2, ISOFORM 1.

```

FT CHAIN 26 258 SRC-LIKE-ADAPTER 2, ISOFORM 2.
FT INIT MET 26 26 FOR ISOFORM 2.
FT DOMAIN 30 90 SH3.
FT DOMAIN 92 189 SH2.
FT DOMAIN 189 258 SLA C-TERMINAL.
FT LIPID 1 1 N-myristoyl glycine.
FT MUTAGEN 1 1 G->A: ABOLISHES LOCALIZATION TO
FT MUTAGEN 1 1 MEMBRANES.
FT MUTAGEN 26 26 M->V: ABOLISHES ISOFORM 2.
FT MUTAGEN 81 81 P->L: DOES NOT AFFECT ITS INHIBITORY
FT FUNCTION.
FT MUTAGEN 119 119 R->E: ABOLISHES INTERACTION WITH ZAP70,
FT AND ITS INHIBITORY FUNCTION.
FT CONFLICT 127 127 C -> Y (IN REF. 3; BAC27168).
FT CONFLICT 159 159 S -> T (IN REF. 3; BAC27168).
FT CONFLICT 160 160 P -> H (IN REF. 2).
FT SEQUENCE 258 AA; 28345 MW; 8270E93DE3FC696A CRC64;

Query Match 75.6%; Score 1023; DB 1; Length 258;
Best Local Similarity 79.3%; Pred. No. 1.8e-78;
Matches 207; Conservative 16; Mismatches 34; Indels 4; Gaps 3;

QY 2 GSLPSRKSLPSLSASSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGELPTI 61
Db 1 GSLSSRKLT-SFSPSSSGPDPQPSVQPERHKVTAVALGSPAGGPAELSLRLGELPTI 59
QY 62 VSEGDGDMWTYLSVSEVSGREYNIPSVHGVKSHGWLVEGLSREKAEELLPLGNPGGAFLIR 121
Db 60 ISEGDGDMWTYQSEVSGREYHMPVYVAKVAHGWLYEGLSREKAEELLPLGNPGGAFLIR 119
QY 122 ESQTRRSYLSVLSRSPASWDRIHRYIHCLDNGWLISPRITFPSLOALVDHYSELAD 181
Db 120 ESQTRRCYLSVLSRSPASWDRIHRYIQRDNGWLISPRITFPSLHALVEHSELAD 179
QY 182 DICLLKEPCVLQAGLPKQIPLPVTQVTPPLNWKELDSSLFSEA-ATGEESLLSFG 240
Db 180 GICCLPEPCVLQGLPGKQPPVPTVPTSLNKKLDRSLFLFLFLEAPASGEASLLSFG 239
QY 241 LRESLSFYISLNDFAVSLDAA 261
Db 240 LRESLSFYISLAED--PLDDA 258

RESULT 3
SLAL_RAT SLAL_RAT STANDARD; PRT; 312 AA.
AC PS3622;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SRC-like-adaptor (Src-like-adaptor protein 1).
GN SLA OR SLAP OR SLAP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Alifragis P., Molnar Z., Parnavelas J.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adaptor protein, which negatively regulates T-cell
CC receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced
CC activation of nuclear factor of activated T-cells. Involved in the
CC negative regulation of positive selection and mitosis of T-cells.
CC May act by linking signaling proteins such as ZAP70 with CBL,
CC leading to a CBL dependent degradation of signaling proteins (By
CC similarity).
CC -!- SUBUNIT: Homodimer. Interacts with phosphorylated CBL, SYK and
CC LAT. Homodimerization and interaction with phosphorylated CBL
CC occurs via its C-terminal domain. Interacts with
CC PDGFRB and EphA2. Interacts with phosphorylated proteins ZAP70;
CC CD32; VAV1 and LCP2 via its SH2 domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; colocalizes with endosomes (By

```

```

similarity).
-!- DOMAIN: The C-terminal domain is essential for the
homodimerization and the interaction with CBL. While the
interaction with CBL is apparently mediated via the hydrophobic
region of this domain, the highly charged region is apparently
required for the homodimerization (By similarity).
-!- PTM: Phosphorylated (By similarity).
-!- SIMILARITY: Contains 1 SH2 domain.
-!- SIMILARITY: Contains 1 SH3 domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY217759; AAC61134.1; -
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001452; SH3.
CC Pfam; PF00017; SH2; 1.
CC PRINTS; PR00018; SH3; 1.
CC PRODOM; PD000093; SH2; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS50001; SH2; 1.
CC PROSITE; PS50002; SH3; 1.
CC SH2 domain; SH3 domain; Phosphorylation.
KW DOMAIN 38 98 SH3.
FT DOMAIN 100 191 SH2.
FT DOMAIN 206 312 SLA C-TERMINAL.
FT SEQUENCE 312 AA; 34744 MW; A05412D39AE68388 CRC64;

Query Match 37.0%; Score 501; DB 1; Length 312;
Best Local Similarity 41.3%; Pred. No. 1.1e-34;
Matches 112; Conservative 41; Mismatches 90; Indels 28; Gaps 6;

QY 11 LPSPLSSSSVQGG-----PVTMEARSKATAV-----ALGSFPAGGPAELSLRLGEP 58
Db 5 LFGFSTRGKEMGNMKSSTAPLERPLNTEGLSDPLAVLNDYPSDIPFPKRGK 64
QY 59 LTVSEGDGDMWTYLSVSEVSGREYNIPSVHGVKSHGWLVEGLSREKAEELLPLGNPGGAFLIR 118
Db 65 LRVISDEGGWKALISLSTGRESYIPGICVARVYHGWLFEGGLGRDKABELLQLPPTKIGSF 124
QY 119 LIRESQTRRSYLSVLSRSPASWDRIHRYIHCLDNGWLISPRITFPSLOALVDHYSE 178
Db 125 MIRESETKKGFYSLSVR-----HRQVXHYIFRLPNNWYIISPRITFQCLEDLVTHYSE 178
QY 179 LADDICLLKKEPCVLQ-----AGLPKQIPLPVTQVTPPLNWKELDSSLFSEAATG- 232
Db 179 VADGLCCVLTTPLAQNTPAPTAPSPCTSPGSPVTLQKTFDKWRVSSLOEGPEGAENP 238
QY 233 ---ESLSLSEGLRESLSFYISL-NDEAVSLD 259
Db 239 LRVEDLSFYSLRESIASYLSLTGDDSNFD 269

RESULT 4
SLAL_MOUSE SLAL_MOUSE STANDARD; PRT; 280 AA.
AC Q60838; Q8C9Q8; Q8CAT0; Q8CBE9; Q8QZX8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SRC-like-adaptor (Src-like-adaptor protein 1) (mSLAP).
GN SLA OR SLAP OR SLAP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

RN [1] SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND INTERACTION  
 RP WITH EPHA2.  
 RC TISSUE=Embryonic brain;  
 RX MEDLINE=95370243; PubMed=7543998;  
 RA Pandey A., Duan H., Dixit V.M.;  
 RT "Characterization of a novel Src-like adapter protein that associates  
 RT with the Src receptor tyrosine kinase.";  
 RL J. Biol. Chem. 270:19201-19204(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=20093985; PubMed=10630289;  
 RA Carrier A., Nguyen C., Victorero G., Granjeaud S., Rocha D.,  
 RA Bernard K., Miazek A., Perrier P., Mallissen M., Nequet P.,  
 RA Mallissen B., Jordan B.R.;  
 RT "Differential gene expression in CD3epsilon- and RAG1-deficient  
 RT thymuses: definition of a set of genes potentially involved in  
 RT thymocyte maturation.";  
 RL Immunogenetics 50:255-270(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=ILS, and ISS;  
 RX MEDLINE=21363810; PubMed=11471062;  
 RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,  
 RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;  
 RT "High-throughput sequence identification of gene coding variants  
 RT within alcohol-related QTLs.";  
 RL Mamm. Genome 12:657-663(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nishida I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Kasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Balcarelli R., Hill D.P., Sult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Guscinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Kongsava A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie I., Miki H.,  
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Akakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Iehii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Smalios D.E.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP INTERACTION WITH PDGFRB.  
 RX MEDLINE=98414809; PubMed=9742401;  
 RA Roche S., Alonso G., Kaziauskas A., Dixit V.M., Courtneidge S.A.,  
 RA Pandey A.;  
 RT "Src-like adaptor protein (Slap) is a negative regulator of  
 RT mitogenesis.";  
 RL Curr. Biol. 8:975-978(1998).  
 RN [7]  
 RP CHARACTERIZATION, FUNCTION, INTERACTION WITH ZAP70; CD3Z; VAV1 AND  
 LCP2, AND MUTAGENESIS OF PRO-72 AND ARG-110.  
 RX MEDLINE=20130290; PubMed=10662792;  
 RA Sosinowski T., Pandey A., Dixit V.M., Weiss A.;  
 RT "Src-like adaptor protein (SLAP) is a negative regulator of T cell  
 RT receptor signaling.";  
 RL J. Exp. Med. 191:463-474(2000).  
 RN [8]  
 RP FUNCTION, SUBCELLULAR LOCALIZATION, MYRISTOYLATION, AND MUTAGENESIS  
 OF GLY-1.  
 RX MEDLINE=20242033; PubMed=10779329;  
 RA Manes G., Bello P., Roche S.;  
 RT "Slap negatively regulates Src mitogenic function but does not revert  
 RT Src-induced cell morphology changes.";  
 RL Mol. Cell. Biol. 20:3396-3406(2000).  
 RN [9]  
 RP FUNCTION, AND DEVELOPMENTAL STAGE.  
 RX MEDLINE=21453118; PubMed=11567635;  
 RA Sosinowski T., Killen N., Weiss A.;  
 RT "The Src-like adaptor protein downregulates the T cell receptor on  
 RT CD4+CD8+ thymocytes and regulates positive selection.";  
 RL Immunity 15:457-466(2001).  
 CC -1- FUNCTION: Adapter protein, which negatively regulates T-cell  
 CC receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced  
 CC activation of nuclear factor of activated T-cells. Involved in the  
 CC negative regulation of positive selection and mitosis of T-cells.  
 CC May act by linking signaling proteins such as ZAP70 with CBL,  
 CC leading to a CBL dependent degradation of signaling proteins.  
 CC -1- SUBUNIT: Homodimer. Interacts with phosphorylated CBL, SYK and  
 CC LAR. Homodimerization and interaction with phosphorylated CBL  
 CC occurs via its C-terminal domain (by similarity). Interacts with  
 CC PDGFRB and EPHA2. Interacts with phosphorylated proteins ZAP70;  
 CC CD3Z; VAV1 and LCP2 via its SH2 domain.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; colocalizes with endosomes.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Name=1;  
 CC IsoId=Q60898-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q60898-2; Sequence=VSP\_007239;  
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in lymphoid tissues.  
 CC Highly expressed in spleen, thymus and lymph nodes. Weakly  
 CC expressed in lung and brain. Expressed in T-cells and at low level  
 CC in B-cells.  
 CC -1- DEVELOPMENTAL STAGE: Expressed during thymocyte maturation. Weakly  
 CC expressed in CD4(-) CD8(-) thymocytes, strongly expressed in  
 CC CD4(+) CD8(+) thymocytes, while expression decreases in more  
 CC mature cells.  
 CC -1- DOMAIN: The C-terminal domain is essential for the  
 CC homodimerization and the interaction with CBL. While the  
 CC interaction with CBL is apparently mediated via the hydrophobic  
 CC region of this domain, the highly charged region is apparently

CC required for the homodimerization (By similarity).

CC -1- PTM: Phosphorylated (By similarity).

CC -1- MISCELLANEOUS: SLA deficient mice show a strong upregulation

CC of TCR and CD5 at the CD4(+) CD8(-) stage, and an enhanced

CC positive selection in T-cells.

CC -1- SIMILARITY: Contains 1 SH2 domain.

CC -1- SIMILARITY: Contains 1 SH3 domain.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U29056; AAA82756.1; -

DR EMBL; AJ131777; CAB66139.1; -

DR EMBL; AY079449; AAL87537.1; -

DR EMBL; AY079450; AAL87538.1; -

DR EMBL; AK036167; BAC29328.1; -

DR EMBL; AK037901; BAC29896.1; -

DR EMBL; AK041565; BAC30988.1; -

DR EMBL; BC032922; AAH32922.1; -

DR HSP; P18277; IBLK.

DR MGD; MG1104295; Sla.

DR InterPro; IPR000980; SH2.

DR InterPro; IPR001452; SH3.

DR Pfam; PF00017; SH2; 1.

DR Pfam; PF00018; SH3; 1.

DR PRINTS; PR00401; SH2DOMAIN.

DR PRODOM; PD000093; SH2; 1.

DR SMART; SM00326; SH3; 1.

DR SMART; SM00326; SH3; 1.

DR PROSITE; PSS0001; SH2; 1.

DR PROSITE; PSS0002; SH3; 1.

DR KW SH2 domain; SH3 domain; Myristate; Phosphorylation;

DR KW Alternative splicing; Lipoprotein.

FT INIT MET 0

FT DOMAIN 21 81 SH3.

FT DOMAIN 83 174 SH2.

FT DOMAIN 189 280 SLA C-TERMINAL.

FT LIPID 1 N-myristoyl glycine.

Query Match 36.3%; Score 491.5; DB 1; Length 280;

Best Local Similarity 43.6%; Pred. No. 6.2e-34;

Matches 115; Conservative 38; Mismatches 82; Indels 29; Gaps 7;

Qy 9 KSLPSPS---LSSSVQGGPVTWEAERSKATVALGSPFAGGPAELSLRLGEPLTIVSED 65

Db 5 KSTSPSPERPLSSS-----EGLESDFLAV-LTDYSPDIPSPPIFRGEXLRVISDE 54

Qy 66 GDWTVTLSEVSGREYNIPSVHVGKVSHGWLVEGLSREKAEELLPLGNPGGAFIRESOT 125

Db 55 GGMKALSLTGRSYIPGICVARVYHGMVFEGLGRDXAEELLQLPDPTKIGSPWIRESET 114

Qy 126 RRGYSYSLVRLSPASPDWRIRHVIHCLDNGWLYISPLRTFPLSLQALVDHYSLADIDIC 185

Db 115 KKGFYSLSVR-----HROVKHVFIFRLPNNWYIISPLRTFQCLDILVTHYSEVADGLCC 168

Qy 196 LLKEPCVLQR-----AGPLPGKDIPLVTVQRTPLNKKEDSSLLPSEAATG-----EESL 236

Db 169 VLTTPICLAQNIPATSPSPCTSPGSPVTRQKTFDWRKVSRLQEGSEGAEPLRVDESIL 228

Qy 237 LSEGLRESLSFYIEL-NDEAVSLD 259

Db 229 PSYGLRESIASYLSLTGDDSSFD 252

RESULT 5

SLA1 HUMAN STANDARD; PR7; 275 AA.

ID SLA1\_HUMAN

AC Q13239; Q9UMQ8;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE SRC-like-adaptor (Src-like-adaptor protein 1) (hSLAP).

GN SLA OR SLAP OR SLAP1.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

[1]

RN SEQUENCE FROM N.A. PubMed=8825655;

RX MEDLINE=9643054;

RA Angrist M., Wells D.E., Chakravarti A., Pandey A.;

RT "Chromosomal localization of the mouse Src-like adaptor protein (Slap)

RT gene and its putative human homolog SLA.";

RL Genomics 30:623-625(1995).

[2]

RN SEQUENCE FROM N.A. AND INDUCTION BY ATRA.

RC TISSUE=Histiocytic lymphoma;

RX MEDLINE=9714857; PubMed=9020066;

RA Ohtsuki T., Hatake K., Ikeda M., Tomizuka H., Terui Y., Uwai M.,

RA Miura Y.;

RT "Expression of Src-like adaptor protein mRNA is induced by all-trans

RT retinoic acid.";

RL Biochem. Biophys. Res. Commun. 230:81-84(1997).

[3]

RN SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.

RC TISSUE=Fetal brain;

RX MEDLINE=98321620; PubMed=9660183;

RA Meijerink P.H.S., Yanakiev P., Zorn I., Grierson A.J., Bikker H.,

RA Dye D., Kalaydjieva L., Baas F.;

RT "The gene for the human Src-like adaptor protein (hSLAP) is located

RT within the 64-kb intron of the thyroglobulin gene.";

RL Eur. J. Biochem. 254:297-303(1998).

[4]

RN SEQUENCE FROM N.A. AND FUNCTION.

RX MEDLINE=21553259; PubMed=11696592;

RA Holland S.J., Liao X.C., Mendenhall M.K., Zhou X., Pardo J., Chu P.,

RA Spencer C., Fu A.C., Sheng N., Yu P., Pali E., Nagin A., Shen M.,

RA Yu S., Chan E., Wu X., Li C., Wolsettschlag M., Aversa G.,

RA Kolbinger F., Bennett M.K., Molineaux S., Luo Y., Payan D.G.,

RA Mancebo H.S.Y., Wu J.;

RT "Functional cloning of Src-like adaptor protein-2 (SLAP-2), a novel

RT inhibitor of antigen receptor signaling.";

RL J. Exp. Med. 194:1263-1276(2001).

[5]

RN SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=22389257; PubMed=12477932;

RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haeh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalón D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Richards S., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[6]

RN SEQUENCE OF 1-71 FROM N.A.

RX MEDLINE=21100465; PubMed=11179692;

RA Kratchmarova I., Sosinowski T., Weiss A., Witter K., Vincenz C.,

RA Pandey A.;  
 RT "Characterization of promoter region and genomic structure of the  
 RL murine and human genes encoding Src like adapter protein.";  
 RL Gene 262:267-273(2001).  
 [7]  
 RN FUNCTION, HOMODIMERIZATION, PHOSPHORYLATION, INTERACTION WITH CBL;  
 RP ZAP70; CD32; SYK AND LAT, AND MUTAGENESIS OF ARG-110; LEU-217;  
 RP LEU-223; LEU-228 AND 236-LEU-238.  
 RX MEDLINE=99380595; PubMed=10449770;  
 RA Tang J., Sawasdikosol S., Chang J.-H., Burakoff S.J.;  
 RT "SLAP, a dimeric adapter protein, plays a functional role in T cell  
 RT receptor signaling";  
 RL Proc Natl Acad Sci U.S.A. 96:9775-9780(1999).  
 CC -!- FUNCTION: Adapter protein, which negatively regulates T-cell  
 CC receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced  
 CC activation of nuclear factor of activated T-cells. Involved in the  
 CC negative regulation of positive selection and mitosis of T-cells.  
 CC May act by linking signaling proteins such as ZAP70 with CBL,  
 CC leading to a CBL dependent degradation of signaling proteins.  
 CC -!- SUBUNIT: Interacts with EPHA2, VAV1, LCP2 and PDGFRB (By  
 CC similarity). Homodimer. Homodimerization and interaction with  
 CC phosphorylated CBL occurs via its C-terminal domain. Interacts  
 CC with phosphorylated proteins ZAP70, CD32, SYK and LAT via its SH2  
 CC domain.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; colocalizes with endosomes (By  
 CC similarity).  
 CC -!- TISSUE SPECIFICITY: Expressed in lung and fetal brain. Weakly  
 CC expressed in heart, adult brain, placenta, liver, skeletal muscle,  
 CC kidney and pancreas.  
 CC -!- INDUCTION: By all-trans retinoic acid (ATRA). Induction is  
 CC indirect and is mediated through other proteins.  
 CC -!- DOMAIN: The C-terminal domain is essential for the  
 CC homodimerization and the interaction with CBL. While the  
 CC interaction with CBL is apparently mediated via the hydrophobic  
 CC region of this domain, the highly charged region is apparently  
 CC required for the homodimerization.  
 CC -!- PTM: Phosphorylated.  
 CC -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U30473; AAC50357.1; -;  
 DR EMBL; D89077; BAA13758.1; -;  
 DR EMBL; U44403; AAC27662.1; -;  
 DR EMBL; BC007042; AAH07042.1; -;  
 DR EMBL; AJ238591; CAB53536.1; -;  
 DR HSSP; P08631; 3HCK.  
 DR Genew; HGNC:10902; SLA.  
 DR MIM; 601099; -;  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR ProDom; PD000093; SH2; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR SH2 domain; SH3 domain; Myristate; Phosphorylation; Lipoprotein.  
 KW INIT MET 0 0  
 BY SIMILARITY.  
 FT DOMAIN 21 81  
 FT SH2.  
 FT SH3.  
 FT SLA C-TERMINAL.  
 FT N-myristoyl glycine (By similarity).  
 FT LIPID 1 1  
 FT R->K: STRONGLY REDUCES INTERACTION WITH  
 FT MUTAGEN 110 110

FT MUTAGEN 217 217 ZAP70, CD32, SYK AND LAT.  
 FT L->S: ABOLISHES INTERACTION WITH CBL,  
 FT WHILE IT DOES NOT AFFECT DIMERIZATION  
 FT WHEN ASSOCIATED WITH S-224 AND S-229.  
 FT L->S: ABOLISHES INTERACTION WITH CBL,  
 FT WHILE IT DOES NOT AFFECT DIMERIZATION  
 FT WHEN ASSOCIATED WITH S-218 AND S-229.  
 FT L->S: ABOLISHES INTERACTION WITH CBL,  
 FT WHILE IT DOES NOT AFFECT DIMERIZATION  
 FT WHEN ASSOCIATED WITH S-218 AND S-224.  
 FT LSL->QSO: ABOLISHES INTERACTION WITH CBL.  
 FT SLIGHTLY AFFECTS DIMERIZATION.  
 FT CONFLICT 70 70 Y -> D (IN REF. 6). CRC64;  
 SQ SEQUENCE 275 AA; 31025 MW; ED6837C99AD534F CRC64;  
 Query Match 35.6%; Score 481.5; DB 1; Length 275;  
 Best Local Similarity 40.3%; Pred. No. 4.1e-33;  
 Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;  
 QY 9 KSLPSPSSSSVQGGPVTMAERSKATAVALGSPAGPAELSLRLGEPLTIYSEDGDM 68  
 DB 5 KSTPAPA-----BRPLNPEGLSDFLAVLSDPSPDISPPFRGKLRVISDEGWM 57  
 QY 69 WTVLSEVSGREYNIPSHVHGKLYEGLSRKAEEILLPGNPGGAFLIRSEQTRRG 128  
 DB 58 WKAISSLSTGRESYIFGICVARVYHGWLFEGLDKABELLQLPDKVGSFWIRESETKG 117  
 QY 129 SYSLSVRLSRPASWDRIHRIHCLDNGWLYISPLTTPSLQALVDHYSELADICLLK 188  
 DB 118 FYLSVSR-----HRQVKHYRIFRLPNWYISPLTTPQLLEDLVNHYSEVADGLCCVLT 171  
 QY 189 EPCVLQRAGPLGKDIPLPVTQVPLNWKELDSSLFSEAATG-----EESLLSRL 241  
 DB 172 TPCLTQSTAAAPRAVSSSPVLRQKTVDRVSR---LQEDPEGTENPLGVDESLSFYGL 228  
 QY 242 RESLSFYSLNDE 254  
 DB 229 RESIASYLSLTSE 241  
 RESULT 6  
 LYN\_HUMAN  
 ID LYN\_HUMAN STANDARD; PRT; 511 AA.  
 AC P07948;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Tyrosine-protein kinase LYN (EC 2.7.1.112).  
 GN LYN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87172710; PubMed=3561390;  
 RA Yamanashi Y., Fukushige S.-I., Samba K., Sukegawa J., Miyajima N.,  
 RA Matsubara K.-I., Yamamoto T., Toyoshima K.;  
 RT "The yes-related cellular gene lyn encodes a possible tyrosine kinase  
 RT similar to p56lck.";  
 RL Mol. Cell. Biol. 7:237-243(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94171041; PubMed=8125304;  
 RA Rider L.G., Raben N., Miller L., Jelsema C.;  
 RT "The cDNAs encoding two forms of the LYN protein tyrosine kinase are  
 RT expressed in rat mast cells and human myeloid cells.";  
 RL Gene 138:219-222(1994).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=LYN A;



CC Name=LYN B; IsoId=P07948-1; Sequence=Displayed;  
 CC IsoId=P07948-2; Sequence=VSP\_005002;  
 CC SIMILARITY: Belongs to the Tyr family of protein kinases. SEC  
 CC subfamily.  
 CC -1- SIMILARITY: Contains 1 SH2 domain.  
 CC -1- SIMILARITY: Contains 1 SH3 domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M16038; AA95540.1; -;  
 CC EMBL; M79321; AA850019.1; -;  
 CC PIR; A26719; TVHULY.  
 CC HSP; P08631; LAD5.  
 CC Genew; HGNC:6735; LYN.  
 CC MIM; 165120; -;  
 CC GO; GO:0004716; F-receptor signaling protein tyrosine kinase . . . ; TAS.  
 CC GO; GO:0006468; P-protein amino acid phosphorylation; TAS.  
 CC GO; GO:0007165; P-signal transduction; TAS.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR000980; SH2.  
 CC InterPro; IPR001452; SH3.  
 CC InterPro; IPR001245; Tyr\_kinase.  
 CC InterPro; IPR008266; Tyr\_kinase\_AS.  
 CC Pfam; PF00069; Pkinase; 1.  
 CC Pfam; PF00017; SH2; 1.  
 CC Pfam; PF00018; SH3; 1.  
 CC PRINTS; PR00401; SH2DOMAIN.  
 CC PRINTS; PR00452; SH3DOMAIN.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC ProDom; PD000093; SH2; 1.  
 CC ProDom; PD000066; SH3; 1.  
 CC SMART; SM00252; SH2; 1.  
 CC SMART; SM00326; SH3; 1.  
 CC SMART; SM00219; TyKc; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE; PS50001; SH2; 1.  
 CC PROSITE; PS50002; SH3; 1.  
 CC Proto-oncogene; Tyrosine-protein kinase; Phosphorylation;  
 CC Transferase; ATP-binding; Myristate; SH2 domain; SH3 domain;  
 CC Palmitate; Lipoprotein; Alternative splicing.  
 CC INT\_MET 0 0  
 CC LIPID 1 1 N-myristoyl glycine (By similarity).  
 CC DOMAIN 2 2 S-palmitoyl cysteine (By similarity).  
 CC DOMAIN 62 122 SH3.  
 CC DOMAIN 128 225 SH2.  
 CC DOMAIN 246 500 PROTEIN\_KINASE.  
 CC NP\_BIND 252 260 ATP (By similarity).  
 CC BINDING 274 274 ATP (By similarity).  
 CC ACT\_SITE 366 366 BY SIMILARITY.  
 CC MOD\_RES 396 396 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 CC MOD\_RES 507 507 PHOSPHORYLATION (BY SIMILARITY).  
 CC VARSPIC 22 42 Missing (in isoform LYN B).  
 CC SEQUENCE 511 AA; 58442 MW; 8419CD461204E364 CRC64;  
 Query Match 27.4%; Score 370.5; DB 1; Length 511;  
 Best Local Similarity 39.8%; Pred No. 1.8e-23;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
 6 SRRKSLPSPSSVQGGPVTMAERSKATPAVALGFPAGPAELSLRIGEPITIVSED 65  
 37 SNKQRPVPE-SQLLPGRFQTKDPEQGDIVWALPYDGIHPDGLSFKKGRMKVLEH 95

QY 66 GDAWTVLSEVSGREYNIPSVHVGK-----SHGWLVEGLSREKAEELLILPGNPGGAFLIR 121  
 Db 96 GEWNAKSLTKKEGIPSNYAKLNTLETESWFFKDIIRKQALAPGNSAGAFLIR 155  
 QY 122 ESOTRGRGYSLSVRLSRPDSWDIRYRHCLDNGWLYISPLTFFSLQALVDHYSELAD 181  
 Db 156 ESETLKGSFSLVRDPPVHGDVIRKIRSLDNGGYISPRITFCISDMIKHYQKQAD 215  
 QY 182 DICLLKPCVLRAGPLPGK 202  
 Db 216 GLCRLEKACI-----SPKPK 232  
 RESULT 7  
 LYN\_MOUSE STANDARD; PRT; 511 AA.  
 AC P25911; Q62127;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Tyrosine-protein kinase LYN (EC 2.7.1.112).  
 GN LYN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91260688; PubMed=1710766;  
 RA Stanley E., Ralph S.J., McEwen S., Boulet I., Holtzman D.A.,  
 RA Lock P., Dunn A.R.;  
 RA "Alternatively spliced murine lym mRNAs encode distinct proteins";  
 RA Mol. Cell. Biol. 11:3399-3406(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91203857; PubMed=2017160;  
 RA Yi T., Bolen J.B., Ihle J.N.;  
 RA "Hematopoietic cells express two forms of lym kinase differing by 21  
 RT amino acids in the amino terminus";  
 RL Mol. Cell. Biol. 11:2391-2398(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM LYN A).  
 RC STRAIN=Czech II; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Schuler G.D.,  
 RA Klausner R.D., Collins P.S., Wagner K.H., Schaefer C.P., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Max S.I., Wang J., Hsieh F.,  
 RA Hopkins R.F., Jordan H., Moore T., Rubin G.M., Hong L.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Casavant T.L., Scheetz T.E.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Carninci P., Prange C.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Mullahy S.J.,  
 RA Raha S.S., Locquellano N.A., Peters G.J., Abramson R.D., Gunaratne P.H.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gay L.J., Hulyk S.W.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Lu X., Gibbs R.A.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Sodergren S., Sanchez A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Green G.G.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 363-431 FROM N.A.  
 RX MEDLINE=90152381; PubMed=2482828;  
 RA Wilks A.F., Kurban R.R., Hovens C.M., Ralph S.J.;  
 RA "The application of the polymerase chain reaction to cloning members  
 RT of the protein tyrosine kinase family";  
 RL Gene 85:167-174(1989).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.



This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL; L14951; AAA41549.1; --  
EMBL; L14782; AAA20944.1; --  
DR EMBL; L14823; AAA20945.1; --  
DR EMBL; AF000300; AAB71344.1; --  
DR EMBL; AF000301; AAB71345.1; --  
DR EMBL; AF000302; AAB71346.1; --  
DR PIR; I56360; I56160.  
DR HSSP; P08631; IAD5.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00069; Ekinase; 1.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR ProDom; PD000093; SH2; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50001; SH2; 1.  
DR PROSITE; PS50002; SH3; 1.  
KW Proto-oncogene; Tyrosine-protein kinase; Phosphorylation;  
KW Transferase; ATP-binding; Myristate; SH2 domain; SH3 domain;  
KW Lipid; Lipoprotein; Alternative splicing.  
FT INIT MET 0  
FT LIPID 1 1 N-myristoyl glycine (By similarity).  
FT LIPID 2 2 S-palmitoyl cysteine (By similarity).  
FT DOMAIN 62 122 SH3.  
FT DOMAIN 128 225 SH2.  
FT DOMAIN 246 500 PROTEIN KINASE.  
FT NP\_BIND 252 260 ATP (By similarity).  
FT BINDING 274 274 ATP (By similarity).  
FT ACT\_SITE 366 366 BY SIMILARITY.  
FT MOD\_RES 396 396 PHOSPHORYLATION (AUTO-) (By similarity).  
FT MOD\_RES 507 507 PHOSPHORYLATION (By similarity).  
FT VARSPIC 24 44 Missing (in isoform LYN B).  
FT CONFLICT 230 230 /FTID=VSP\_005004.  
FT CONFLICT 307 307 P -> L (IN REF. 2).  
FT CONFLICT 418 418 V -> A (IN REF. 2).  
FT CONFLICT 511 511 C -> Y (IN REF. 2).  
FT SEQUENCE 511 AA; 58529 MW; 24A2E5E229CD43ED CRC64;  
Query Match 27.2%; Score 367.5; DB 1; Length 511;  
Best Local Similarity 40.3%; Pred. No. 3.2e-23;  
Matches 81; Conservative 33; Mismatches 78; Indels 9; Gaps 3;  
QY 6 SRKSLPSLSSVQGGPVTMEATSKATAVALGSPGPAELSLRLGELPTIVISED 65  
Db 37 SNKQRPYPE-SQLPGQRFQAKDFEQGDIVVALYDGIHPDLSFKKGEKMKVLEEH 95  
QY 66 GDMWTVLSVSGREYNISVHVGVK-----SHGWLVEGLSRKAEBLLLLPQNGAFILR 121  
Db 96 GSWWKAKLSSKREGFIPSNVYKVTLETTEWFEFKDITRKDAERQLLAPGNSAGAFILR 155  
QY 122 ESCRTRGYSLSVLRSPASWDRIRHYRTHCLDNGWLYISPRLTTPPSLQALVDHYSELAD 181

Db 156 ESETLKGSFSLSVRDYPMHGDVHKYKIRSLDNGGYVISPRITPCISDMIKHYQKQSD 215  
QY 182 DICLLKEPCVLQAGPLPGK 202  
Db 216 GLCRRLEKACI-----SPKPK 232  
RESULT 9  
HCK\_HUMAN  
ID HCK\_HUMAN STANDARD; PRT; 525 AA.  
AC P08631; Q96CC0; Q9HSY5; Q9NUA4; Q9UNJ5;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tyrosine-protein kinase HCK (BC 2.7.1.112) (p59-HCK/p60-HCK)  
DE (Hemopoietic cell kinase).  
GN HCK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE OF 21-525 FROM N.A.  
RX MEDLINE=87257942; PubMed=3496523;  
RA Quintrell N., Lebo R., Varmus H., Bishop J.M., Pettenati M.J.,  
le Beau M.M., Diaz M.C., Rowley J.D.; that encodes a protein-tyrosine  
RT "Identification of a human gene (HCK) that encodes a protein-tyrosine  
RT kinase and is expressed in hemopoietic cells."  
RL Mol. Cell. Biol. 7:2267-2275(1987).  
RN [2]  
RP SEQUENCE OF 21-525 FROM N.A.  
RX MEDLINE=87257943; PubMed=3453117;  
RA Ziegler S.F., March J.D., Lewis D.B., Perlmutter R.M.;  
RT "Novel protein-tyrosine kinase gene (hck) preferentially expressed in  
RT cells of hematopoietic origin."  
RL Mol. Cell. Biol. 7:2276-2285(1987).  
RN [3]  
RP SEQUENCE OF 21-525 FROM N.A.  
RC TISSUE-B-cell.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.O., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 21-525 FROM N.A.  
RC TISSUE=ileal mucosa;  
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,  
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,  
RA Ohayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,  
RA Isogai T., Sugano S.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor P.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dunn P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.D., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Leinvaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
RA Stuce C.D., Smith M.L., Soderlund C., Stewart C.A., Suleston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Thorpe A.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RT Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
[6]  
RP SEQUENCE OF 178-525 FROM N.A.  
RX TISSUE=Spleen;  
RC MEDLINE=92241680; PubMed=1572549;  
RA Hradetzky D., Streibhardt K., Ruebsamen-Waigmann H.;  
RT "The genomic locus of the human hemopoietic-specific cell protein  
RT tyrosine kinase (PTK)-encoding gene (HCK) confirms conservation of  
RT exon-intron structure among human PTKs of the scr family.";  
RL Gene 113:275-280(1992).  
[7]  
RP SEQUENCE OF 1-21 FROM N.A., AND ALTERNATIVE INITIATION.  
RX MEDLINE=91342636; PubMed=1875927;  
RA Lock P., Ralph S., Stanley E., Boulet I., Ramsay R., Dunn A.R.;  
RT "Two isoforms of murine hck, generated by utilization of alternative  
RT translational initiation codons, exhibit different patterns of  
RT subcellular localization.";  
RL Mol. Cell. Biol. 11:4363-4370(1991).  
[8]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 77-525.  
RX MEDLINE=97177106; PubMed=9024658;  
RA Sighieri F., Moarefi I., Kuriyan J.;  
RT "Crystal structure of the Src family tyrosine kinase Hck.";  
RL Nature 385:602-609(1997).  
[9]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 80-136.  
RX MEDLINE=98453315; PubMed=9778343;  
RA Arold S., O'Brien R., Franken P., Strub M.P., Hoh F., Dumas C.,  
RA Ladbury J.E.;  
RT "T-loop flexibility enhances the specificity of Src family SH3  
RT domains for HIV-1 Nef.";  
RL Biochemistry 37:14683-14691(1998).  
[10]  
RP STRUCTURE BY NMR OF 77-137.  
RX MEDLINE=98239731; PubMed=9571048;  
RA Horita D.A., Baldisseri D.M., Zhang W., Altieri A.S., Smithgall T.E.,  
RA Gmeiner W.H., Byrd R.A.;  
RT "Solution structure of the human Hck SH3 domain and identification of  
RT its ligand binding site.";  
RL J. Mol. Biol. 278:253-265(1998).  
[11]  
RP STRUCTURE BY NMR OF 138-244.  
RX MEDLINE=97263487; PubMed=9109402;  
RA Zhang W., Smithgall T.E., Gmeiner W.H.;  
RT "Sequential assignment and secondary structure determination for the  
RT Src homology 2 domain of hematopoietic cellular kinase.";  
RL FEBS Lett. 406:131-135(1997).

CC -!- FUNCTION: May serve as part of a signaling pathway coupling the Fc  
CC receptor to the activation of the respiratory burst. May also  
CC contribute to neutrophil migration and may regulate the  
CC degradation process of neutrophils.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Isoform p60-HCK and isoform p59-HCK are  
CC associated with membranes. Isoform p60-HCK is also cytoplasmic (By  
CC similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative initiation;  
CC Comment=2 isoforms, p60-HCK (shown here) and p59-HCK, are  
CC produced by alternative initiation;  
CC -!- TISSUE SPECIFICITY: Expressed predominantly in cells of the  
CC myeloid and b-lymphoid lineages  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M16591; AAA52643.1; -;  
CC EMBL: M16592; AAA52644.1; -;  
CC EMBL: BC014435; AAH14435.1; -;  
CC EMBL: AK026432; BAB15482.1; -;  
CC EMBL: AL049539; CAB75606.1; -;  
CC EMBL: X58741; CAA41565.2; -;  
CC EMBL: X58742; CAA41565.2; JOINED.  
CC EMBL: X58743; CAA41565.2; JOINED.  
CC PIR: A27811; TVHUHC  
CC FDB: 2HCK; 20-AUG-97.  
CC FDB: 3HCK; 15-OCT-97.  
CC FDB: 4HCK; 17-JUN-98.  
CC FDB: 5HCK; 17-JUN-98.  
CC FDB: 1AD5; 15-MAY-97.  
CC FDB: 1BUL; 11-NOV-98.  
CC FDB: 1OCF; 08-JUN-99.  
CC Genew; HGNC:4840; HCK.  
CC MIM: 142370; -;  
CC GO: GO:0004713; F:protein-tyrosine kinase activity; TAS.  
CC GO: GO:0007498; P:mesoderm development; TAS.  
CC GO: GO:0006468; P:protein amino acid phosphorylation; TAS.  
CC InterPro: IPR000719; Prot\_kinase.  
CC InterPro: IPR000980; SH2.  
CC InterPro: IPR001452; SH3.  
CC InterPro: IPR001245; Tyr\_kinase.  
CC InterPro: IPR008266; Tyr\_kinase\_AS.  
CC Pfam: PF00069; pkinase; 1.  
CC Pfam: PF00017; SH2; 1.  
CC Pfam: PF00018; SH3; 1.  
CC PRINTS: PR00401; SH2DOMAIN.  
CC PRINTS: PR00452; SH3DOMAIN.  
CC ProDom: PD000001; Prot\_kinase; 1.  
CC ProDom: PD000093; SH2; 1.  
CC ProDom: PD000066; SH3; 1.  
CC -----

Query Match 26.6%; Score 360.5; DB 1; Length 525;  
Best Local Similarity 41.6%; Pred. No. 1.3e-22;  
Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;

QY 12 PPSLSVVQGGPVTMEERSKATAVALGSPAGGPAELSLRLGELPTIVSGDQWTV 71  
Db 60 PGENSENS---NTPGIREAGSEDIIVVLYDEALHHDLSPFGKQDQWVLEESGEWKA 116  
QY 72 LSEVSGREYNIPSVHVGVKVV---SHGWLYEGLSRKAEELLLPGNPGAGFLIRSQTR 127

DR PROSITE: P550002; SH3: 1.  
KW Transferase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;  
KW Myristate; SH2 domain; SH3 domain; Lipoprotein.  
FT INIT MET 0 0  
FT DOMAIN 55 115 SH3.  
FT DOMAIN 121 218 SH2.  
FT DOMAIN 239 492 PROTEIN KINASE.  
FT NP BIND 245 253 ATP (BY SIMILARITY).  
FT BINDING 267 267 ATP (BY SIMILARITY).  
FT ACT SITE 358 358 ATP (BY SIMILARITY).  
FT LIPID 1 N-myristoyl glycine (By similarity).  
FT MOD\_RES 388 388 PHOSPHORYLATION (AUTO-). (BY SIMILARITY).  
FT SEQUENCE 503 AA; 56964 MW; 561F9322D2DE3436 CRC64;  
Query Match 26.5%; Score 358; DB 1; Length 503;  
Best Local Similarity 45.3%; Pred. NO. 2e-22;  
Matches 72; Conservative 29; Mismatches 54; Indels 4; Gaps 1;  
QY 38 VALGSPAGGPAELSLRLCEPLTIIVSEGDWTVLSEVSGREYNIPSVHVGVK- ---SHG 93  
DB 61 VALDYEAITHHDLSPQKGDQWVLEESGWNKARSLATKRGYISFNVARVDSLETEE 120  
QY 94 WLYEGLSREKAELELLLPNGPGAFILRSQTRRGYSLSVRLSRPASWDRIHYRIHCL 153  
DB 121 WFFKGISRKDAERQLLAPGNMLGSFMRDSETTKGSYLSVRDYPDPRQGDVVKHKIRTL 180  
QY 154 DNGWLVISPLRPFSIQALVDHYSELADDCILLKEPCV 192  
DB 181 DNGGFIISPRSFSTIQELVDHYKSGDGLCQKLSVPCV 219  
RESULT 11  
BLK\_HUMAN STANDARD; PRT; 504 AA.  
ID BLK\_HUMAN AC P51451; Q16291;  
AC 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DT Tyrosine-protein kinase BLK (EC 2.7.1.112) (B lymphocyte kinase) (p55-  
DE BLK).  
DE Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RP MEDLINE=95123078; PubMed=7822795;  
RA Drebin J.A., Hartzell S.W., Griffin C., Campbell M.J.,  
RA Niederhuber J.E.;  
RA "Molecular cloning and chromosomal localization of the human homologue  
RT of a B-lymphocyte specific protein tyrosine kinase (blk).";  
RL Oncogene 10:477-485(1995).  
CC -!- FUNCTION: May function in a signal transduction pathway that is  
CC restricted to B lymphoid cells.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC The European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DE EMBL; Z33998; CAA83965.1; -  
CC EMBL; S76617; AAB33265.1; -  
CC PIR; I37206; I37206.  
DR HSP; P16277; IBLK.  
DR Genew; HGNC:1057; BLK.  
DR MM; 191305; -  
DR GO; GO:0004713; F-protein-tyrosine kinase activity; TAS.  
DR GO; GO:0007243; P-protein kinase cascade; TAS.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR ProDom; PD000093; SH2; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SMO0252; SH2; 1.  
DR SMART; SMO0326; SH3; 1.  
DR SMART; SMO0219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00109; PROTEIN KINASE\_TYR; FALSE\_NEG.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00001; SH2; 1.  
DR PROSITE; PS00002; SH3; 1.  
DR Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding; Myristate; SH2 domain; SH3 domain; Lipoprotein.  
FT INIT\_MET 0 0  
FT LIPID 1 1  
FT DOMAIN 57 117  
FT DOMAIN 123 219  
FT DOMAIN 240 493  
FT NP\_BIND 246 254  
FT BINDING 268 268  
FT ACT\_SITE 359 359  
FT MOD\_RES 388 388  
FT CONFLICT 286 286  
FT CONFLICT 406 406  
SQ SEQUENCE 504 AA; 57607 MW; BDB1DF50EC7370C8 CRC64;  
Query Match 26.1%; Score 352.5; DB 1; Length 504;  
Best Local Similarity 43.6%; Pred. No. 5,7e-22;  
Matches 75; Conservative 24; Mismatches 58; Indels 5; Gaps 2;  
QY 25 PVTWEAKSKATAVALGSPFAGGPAELSLRIGELPTTVSDGDWWTVLVSQGVREYNIPS 84  
Db 50 PPDEHLDKHFVVALDYDTAMNDQLQMLKGEKLVQLKGTGDWTLARSILVTGREGVPS 109  
QY 85 VHVGVKVS---HGWLVEGLSKAEAEILLPGNPGGAFLIREOTRGYSLSVLSRPA 140  
Db 110 NFVARVESLEMERFFRSQGRKEARQLLAPINKAGSFLIRESETNKGAFSLVK-DVTT 168  
QY 141 SWDRIRHYRIHCLDNGMLYISPRITFTPSLQALVDHYSELADDICCLKEPCV 192  
Db 169 QGELIKHYKIRCLDEGGYIISPRITFTPSLQALVQHYSKYKGGGLCQLRLTFCV 220  
RESULT 12  
HCK\_MOUSE ID HCK\_MOUSE  
AC P08103  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Tyrosine-protein kinase HCK (EC 2.7.1.112) (p56-HCK/p59-HCK)  
CC (Hemopoietic cell kinase) (B-cell/myeloid kinase) (BMK).  
CC HCK.  
OS Mus musculus (Mouse).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
CC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE OF 21-523 FROM N.A.  
RC STRAIN=ICR; TISSUE=Macrophage;  
RX MEDLINE=88067781; PubMed=3684607;  
RA Klemz M.J., McKercher S.R., Maki R.A.;  
RT "Nucleotide sequence of the mouse hck gene.";  
RL Nucleic Acids Res. 15:9600-9600(1987).  
RN [2]  
RP SEQUENCE OF 21-523 FROM N.A.  
RX MEDLINE=88068587; PubMed=3317404;  
RA Holtzman D.A., Cook W.D., Dunn A.R.;  
RT "Isolation and sequence of a cDNA corresponding to a src-related gene expressed in murine hemopoietic cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8325-8329(1987).  
RN [3]  
RP SEQUENCE OF 1-21 FROM N.A., AND ALTERNATIVE INITIATION.  
RX MEDLINE=91342636; PubMed=1875927;  
RA Lock P., Ralph S., Stanley E., Boulet J., Ramsay R., Dunn A.R.;  
RT "Two isoforms of murine hck, generated by utilization of alternative translational initiation codons, exhibit different patterns of subcellular localization.";  
RL Mol. Cell. Biol. 11:4363-4370(1991).  
CC !- FUNCTION: May serve as part of a signaling pathway coupling the Fc receptor to the activation of the respiratory burst. May also contribute to neutrophil migration and may regulate the degranulation process of neutrophils.  
CC !- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC !- SUBCELLULAR LOCATION: Isoform p59-HCK and isoform p56-HCK are associated with membranes. Isoform p59-HCK is also cytoplasmic.  
CC !- ALTERNATIVE PRODUCTS:  
CC Event-Alternative initiation;  
CC Comment=2 isoforms, p59-HCK (shown here) and p56-HCK, are produced by alternative initiation;  
CC !- TISSUE SPECIFICITY: Expressed predominantly in cells of the myeloid and b-lymphoid lineages.  
CC !- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC subfamily.  
CC !- SIMILARITY: Contains 1 SH2 domain.  
CC !- SIMILARITY: Contains 1 SH3 domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DE EMBL; Y00487; CAA68544.1; -  
CC EMBL; J03023; AAA37305.1; -  
CC PIR; A27282; TWGSHC.  
DR HSP; P08631; IAD5.  
DR MGD; MG1:96052; HCK.  
DR InterPro; IPR00108; Neu\_cyt\_fact\_2.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00499; P67PHOX.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00452; SH3DOMAIN.

Okano Y., Sugimoto Y., Fukuoka M., Matsui A., Nagata K.I., Nozawa Y.:  
"Identification of rat cDNA encoding hck tyrosine kinase from  
megakaryocytes.";  
Biochem. Biophys. Res. Commun. 181:1137-1144 (1991).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=Wistar; TIGSUS=Spleen;  
Vijaya Gouri B.S., Rema V., Kamatkar S., Swarup G.:  
"Nucleotide sequence of a cDNA coding for rat hck tyrosine kinase and  
characterization of its gene product.";  
J. Biosci. 19:117-129 (1994).  
CC -1- FUNCTION: May serve as part of a signaling pathway coupling the Fc  
receptor to the activation of the respiratory burst. May also  
contribute to neutrophil migration and may regulate the  
degranulation process of neutrophils.  
CC CC CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated.  
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
subfamily.  
CC -1- SIMILARITY: Contains 1 SH2 domain.  
CC -1- SIMILARITY: Contains 1 SH3 domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
EMBL; S74141; AAB20754.1; --  
EMBL; M83666; AAA41312.1; --  
EMBL; X62345; CAA44218.1; --  
PIR; JQ1321; JQ1321.  
DR HSP; P08631; XEUI.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR ProDom; PD000093; SH2; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SMO0252; SH2; 1.  
DR SMART; SMO0326; SH3; 1.  
DR SMART; SMO0219; Tyrcg; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50001; SH2; 1.  
DR PROSITE; PS50002; SH3; 1.  
DR Transferrase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;  
KW Myristate; SH2 domain; SH3 domain; Lipoprotein.  
KW INIT NET 0  
FT DOMAIN 54 114  
FT SH2.  
FT DOMAIN 120 217  
FT PROTEIN KINASE.  
FT DOMAIN 238 491  
FT NP BIND 244 252  
FT ATP (BY SIMILARITY).  
FT BINDING 266 266  
FT ATP (BY SIMILARITY).  
FT ACT SITE 357 357  
FT BY SIMILARITY.  
FT LIPID 1 1  
FT N-myristoyl glycine (By similarity).  
FT MOD RES 387 387  
FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CONFLICT 50 50  
FT CONFLICT 204 204  
FT CONFLICT 305 305  
FT CONFLICT 305 305  
FT CONFLICT 502 AA; 56885 MW; 4CFC1F3F0EB2EADF CRC64;  
SQ SEQUENCE

```
Query Match      26.0%; Score 351.5; DB 1; Length 502;
Best Local Similarity 40.9%; Pred. No. 6.9e-22;
Matches 83; Conservative 32; Mismatches 73; Indels 15; Gaps 4;

QY 5 PSRRKSLPSLSLSSVQGGPVTMEARSKAT-AVALGSPFAGGPAELSLRLGEPITIVS 63
DB 32 PTPSPKLGPSINSLPG-----FVGESEDTIVALDYEAHREDLSFKQGDQMVLE 85
QY 64 EDGQWTVLSEVSGREYNIPSVHVKV-----SHGWLVEGLSRKAEKELLLLPNGPGAPL 119
DB 86 ESGEWKARSLATKEGYIPSNYARVNSLTEWFFKGISRKDAERHLLAPGNLGSFM 145
QY 120 IRESQTRGYSLSVRLSPASMDRIRHYRHICLDNGWLVIYSPRLTPPSLQALVDHYSEL 179
DB 146 IRDSETRKGSVLSVRDFDPQHGDTVXHYKIRTLDSGGFVISPRTFSSIQELVWHYKKG 205
QY 180 ADDICLLKPCVLRAGPLPGK 202
DB 206 KDGLCQKLSVPCV-----SPKPK 224

RESULT 14
BLK_MOUSE
ID BLK_MOUSE STANDARD; PRT; 498 AA.
AC P16277;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase BLK (BC 2.7.1.112) (B lymphocyte kinase) (p55-
DE BLK).
GN BLK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=B-cell;
RC MEDLINE=90117147; PubMed=2404338;
RA Dymecki S.M., Niederhuber J.E., Desiderio S.V.;
RT "Specific expression of a tyrosine kinase gene, blk, in B lymphoid
RT cells.";
RL Science 247:332-336(1990).
RN [2]
RP STRUCTURE BY NMR OF SH2 DOMAIN.
RX MEDLINE=96224819; PubMed=8639560;
RA Metzler W.J., Leitling B., Pryor K., Mueller L., Farmer B.T. II;
RT "The three-dimensional solution structure of the SH2 domain from
RT p55blk kinase.";
RL Biochemistry 35:6201-6211(1996).
CC -!- FUNCTION: May function in a signal transduction pathway that is
CC -!- restricted to B lymphoid cells.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC -!- tyrosine phosphate.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC
CC -!- subfamily.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis-sib.ch/
CC or send an email to license@sib-sib.ch).
```

```
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001454; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000093; SH2; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; FALSE_NEG.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00002; SH3; 1.
KW Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding;
KW Myristate; SH2 domain; SH3 domain; 3D-structure; Lipoprotein.
FT INIT_MET 0 0
FT LIPID 1 1
FT DOMAIN 51 111
FT DOMAIN 117 213
FT DOMAIN 234 487
FT NP_BIND 240 248
FT BINDING 262 262
FT ACT_SITE 353 353
FT MOD_RES 382 382
FT HELIX 124 131
FT TURN 132 132
FT TURN 134 135
FT TURN 138 139
FT STRAND 141 145
FT TURN 147 148
FT STRAND 153 157
FT STRAND 159 161
FT TURN 162 163
FT STRAND 164 166
FT STRAND 170 175
FT TURN 176 178
FT STRAND 179 183
FT TURN 184 185
FT STRAND 186 188
FT HELIX 191 197
FT TURN 198 200
SQ SEQUENCE 498 AA; 56513 MW; BE49D7B079FDD577 CRC64;

Query Match      25.2%; Score 341.5; DB 1; Length 498;
Best Local Similarity 40.0%; Pred. No. 4.7e-21;
Matches 74; Conservative 29; Mismatches 65; Indels 17; Gaps 3;

QY 12 PPSLSVSSVQGGPVTMEARSKATAVAGSPFAGGPAELSLRLGEPITIVSGDWTV 71
DB 43 PPSN-----QDPDEERFVAFVADYAAVNDRLQVLKGEKQLVLRSTGDMWLA 90
QY 72 LSEVSGREYNIPSVHVKV-----HGWLVEGLSRKAEKELLLLPNGPGAFILRESQTR 127
DB 91 RSLVTREGYVPSNFVAPVETLEVKWFFRTSRKDAERQLLAPNMKAGSFLIRSESNK 150
QY 128 GSYLSVRLSPASMDRIRHYRHICLDNGWLVIYSPRLTPPSLQALVDHYSELADICLL 187
DB 151 GAPSLSVK-DITTOGQGVVXHYKIRSLDNGGYVISPRTITLQALVQHYSKKGDLGCL 209

QY 188 KPCV 192
DB 210 TLPCV 214
```



RESULT 15  
 LCK\_HUMAN STANDARD; PRT; 508 AA.  
 AC P06239; P07100; Q12850; Q13152; Q96D04; Q9NYT8;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Proto-oncogene tyrosine-protein kinase LCK (EC 2.7.1.112) (P56-LCK)  
 DE (LCK) (T cell-specific protein-tyrosine kinase).  
 GN LCK  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89123626; PubMed=3265417;  
 RA Perlmutter R.M., Marth J.D., Lewis D.B., Peet R., Ziegler S.F.,  
 RA Wilsson C.B.;  
 RT "Structure and expression of lck transcripts in human lymphoid  
 RT cells";  
 RL J. Cell. Biochem. 38:117-126 (1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87133831; PubMed=3493153;  
 RA Koga Y., Caccia N., Toyonaga B., Spolski R., Yanagi Y., Yoshikai Y.,  
 RA Mak T.W.;  
 RT "A human T cell-specific cDNA clone (vlt6) encodes a protein with  
 RT extensive homology to a family of protein-tyrosine kinases";  
 RL Eur. J. Immunol. 16:1643-1646 (1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90108697; PubMed=2558056;  
 RA Rouer E., van Huynh T., de Souza S.L., Lang M.C., Fischer S.,  
 RA Benarous R.;  
 RT "Structure of the human lck gene: differences in genomic organisation  
 RT within src-related genes affect only N-terminal exons";  
 RL Gene 84:105-113 (1989).  
 RN [4]  
 RP SEQUENCE FROM N.A., VARIANTS LEU-27; GLN-LYS-PRO-231 INS; VAL-352  
 RP AND LEU-446, AND PHOSPHORYLATION OF TYR-393 AND TYR-504.  
 RC TISSUE=Leukemia;  
 RX MEDLINE=94187714; PubMed=9139546;  
 RA Wright D.D., Setton B.W., Kamps M.P.;  
 RT "Oncogenic activation of the lck protein accompanies translocation of  
 RT the LCK gene in the human HSB2 T-cell leukemia";  
 RL Mol. Cell. Biol. 14:2429-2437 (1994).  
 RN [5]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Leukemic T-cell;  
 RX MEDLINE=96085119; PubMed=7495859;  
 RA Vogel L.B., Arthur R., Fujita D.J.;  
 RT "An aberrant lck mRNA in two human T-cell lines";  
 RL Biochim. Biophys. Acta 1264:168-172 (1995).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lymph;  
 RX MEDLINE=23388257; PubMed=12477932;  
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diachenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Matek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Kzywinski M.I., Skalska U., Smalilus D.E.,  
 RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [7]  
 RP SEQUENCE OF 13-508 FROM N.A.  
 RC TISSUE=Peripheral blood lymphocytes;  
 RX MEDLINE=20462621; PubMed=11009057;  
 RA Boncristiano M., Majolini M.B., D'Elia M.M., Pacini S., Valensin S.,  
 RA Olivieri C., Amadei A., Fallini B., Del Prete G., Telford J.L.,  
 RA Baldari C.T.;  
 RT "Defective recruitment and activation of ZAP-70 in common variable  
 RT immunodeficiency patients with T cell defects";  
 RL Eur. J. Immunol. 30:2632-2638 (2000).  
 RN [8]  
 RP SEQUENCE OF 367-508 FROM N.A.  
 RX MEDLINE=89217332; PubMed=2835736;  
 RA Veillette A., Foss F.M., Sausville E.A., Bolen J.B., Rosen N.;  
 RT "Expression of the lck tyrosine kinase gene in human colon carcinoma  
 RT and other non-lymphoid human tumor cell lines";  
 RL Oncogene Res. 1:357-374 (1987).  
 RN [9]  
 RP SEQUENCE OF 374-508 FROM N.A.  
 RX MEDLINE=87000726; PubMed=3489486;  
 RA Trevillian J.M., Lin Y., Chen S.J., Phillips C.A., Canna C.,  
 RA Linna T.J.;  
 RT "Human T lymphocytes express a protein-tyrosine kinase homologous to  
 RT p56LSTRA";  
 RL Biochim. Biophys. Acta 888:286-295 (1986).  
 RN [10]  
 RP SEQUENCE OF 1-34 FROM N.A.  
 RX MEDLINE=89096891; PubMed=2850479;  
 RA Garvin A.M., Pawar S., Marth J.D., Perlmutter R.M.;  
 RT "Structure of the murine lck gene and its rearrangement in a murine  
 RT lymphoma cell line";  
 RL Mol. Cell. Biol. 8:3058-3064 (1988).  
 RN [11]  
 RP SEQUENCE OF 1-34 FROM N.A.  
 RX MEDLINE=89313764; PubMed=2787474;  
 RA Takadera T., Leung S., Gernone A., Koga Y., Takiyama Y.,  
 RA Miyamoto N.G., Mak T.W.;  
 RT "Structure of the two promoters of the human lck gene: differential  
 RT accumulation of two classes of lck transcripts in T cells";  
 RL Mol. Cell. Biol. 9:2173-2180 (1989).  
 RN [12]  
 RP MASS SPECTROMETRY.  
 RC TISSUE=Breast cancer;  
 RX MEDLINE=21829512; PubMed=11840567;  
 RA Harris R.A., Yang A., Stein R.C., Lucy K., Brusten L., Herath A.,  
 RA Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J.,  
 RA Zvelebil M.J.;  
 RT "Cluster analysis of an extensive human breast cancer cell line  
 RT protein expression map database";  
 RL Proteomics 2:212-223 (2002).  
 RN [13]  
 RP INTERACTION WITH PI3K.  
 RX MEDLINE=94067101; PubMed=7504174;  
 RA Vogel L.B., Fujita D.J.;  
 RT "The SH3 domain of p56lck is involved in binding to  
 RT phosphatidylinositol 3'-kinase from T lymphocytes";  
 RL Mol. Cell. Biol. 13:7408-7417 (1993).  
 RN [14]  
 RP INTERACTION WITH KHDRBS1.  
 RX MEDLINE=95155308; PubMed=7852312;  
 RA Vogel L.B., Fujita D.J.;  
 RT "p70 phosphorylation and binding to p56lck is an early event in  
 RT interleukin-2-induced onset of cell cycle progression in  
 RT T-lymphocytes";  
 RL J. Biol. Chem. 270:2506-2511 (1995).  
 RN [15]  
 RP PHOSPHORYLATION OF TYR-504.  
 RX MEDLINE=92347326; PubMed=1639064;

Search completed: February 20, 2004, 09:30:30  
Job time : 41 secs

RA Bergman M., Mustelin T., Oetken C., Partanen J., Flint N.A.,  
RA Amrein K.E., Autero M., Burn P., Altalo K.;  
RT "The human p50csk tyrosine kinase phosphorylates p56lck at Tyr-505 and  
RT down regulates its catalytic activity.";  
RL EMBO J. 11:2919-2924(1992).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 52-225.  
RX MEDLINE=94203291; PubMed=7512222;  
RA Eck M.J., Atweell S.K., Shoelson S.E., Harrison S.C.;  
RT "Structure of the regulatory domains of the Src-family tyrosine  
RT kinase Lck.";  
RL Nature 368:764-769(1994).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 126-220.  
RX MEDLINE=95173978; PubMed=7332720;  
RA Mikol V., Baumann G., Keller T.H., Manning U.M., Zurini M.G.M.;  
RT "The crystal structures of the SH2 domain of p56lck complexed with  
RT two phosphopeptides suggest a gated peptide binding site.";  
RL J. Mol. Biol. 246:344-355(1995).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (1.0 ANGSTROMS) OF 121-225.  
RX MEDLINE=96177765; PubMed=8604142;  
RA Tong L., Warren T.C., King J., Betageri R., Rose J., Jakes S.;  
RT "Crystal structures of the human p56lck SH2 domain in complex with  
RT two short phosphotyrosyl peptides at 1.0-A and 1.8-A resolution.";  
RL J. Mol. Biol. 256:601-610(1996).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 118-225.  
RX MEDLINE=98352059; PubMed=9685372;  
RA Tong L., Warren T.C., Lukas S., Schembri-King J., Betageri R.,  
RA Proudfoot J.R., Jakes S.;  
RT "Carboxymethyl-phenylalanine as a replacement for phosphotyrosine in  
RT SH2 domain binding.";  
RL J. Biol. Chem. 273:20238-20242(1998).  
RN [20]  
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 230-500.  
RX MEDLINE=97100952; PubMed=8945479;  
RA Yamaguchi H., Hendrickson W.A.;  
RT "Structural basis for activation of human lymphocyte kinase Lck upon  
RT tyrosine phosphorylation.";  
RL Nature 384:484-489(1996).  
CC -!- FUNCTION: May participate in antigen-induced T-cell activation.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBUNIT: Binds to phosphatidylinositol 3'-kinase (PI3K) from T  
CC lymphocytes through its SH3 domain and to the tyrosine  
CC phosphorylated form of KDRBS1/p70 through its SH2 domain.  
CC -!- SUBCELLULAR LOCATION: BOUND TO THE CYTOPLASMIC DOMAIN OF EITHER  
CC CD4 OR CD8.  
CC -!- ALTERNATIVE PRODUCTS.  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long;  
CC IsoId=P06239-1; Sequences=Displayed;  
CC Name=Short;  
CC IsoId=P06239-2; Sequences=VSP 005000, VSP 005001;  
CC Note=No experimental confirmation available;  
CC  
Query Match 25.1%; Score 340; DB 1; Length 508;  
Best Local Similarity 40.6%; Pred. No. 6.4e-21;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
QY 26 VTMEAKRSKAT-----AVALGSPAGGPAELSLRLEPLATVSDGDWTVLSEVSGRE 79  
DB 48 VTVEGSPASPQDNVLVIALHSEYPSHDGLDPEKGEQLRIEQSGEWKQAQSLTTGQE 107  
QY 80 YNIPSVHGVKVS-----HGWLYEGLGREKAEELLILPGNPGGAFLIREQTRGYSLSVR 135  
DB 108 GRIPFNFAKANSLEPEPWFNFKLSRKABRQLAPGNTHGSLFIRESESTAGSFSLSVR 167  
QY 136 LSRPASWDRIRVRIHCLDNGWLYISPRLTTPSLQALVDHYSELADDCCLIKPCVQLQR 195  
DB 168 DFDQNGEVVVKYKTNLDNGGFIISPRITFFGLHVLVRYHTNASDGLCTLSRQCQTK 227

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:27:57 ; Search time 45 Seconds  
(without alignments)  
299.431 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGS1PSRRKSLPSPSLSSV.....RESLSFVSLNDAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	370.5	27.4	512	4	US-08-426-509A-16
2	370.5	27.4	512	4	US-08-232-545-16
3	370.5	27.4	512	5	PCT-US95-05008-16
4	360.5	26.6	505	4	US-08-426-509A-17
5	360.5	26.6	505	4	US-08-232-545-17
6	360.5	26.6	505	5	PCT-US95-05008-17
7	344.5	25.5	499	4	US-08-426-509A-19
8	344.5	25.5	499	4	US-08-232-545-19
9	344.5	25.5	499	5	PCT-US95-05008-19
10	340	25.1	508	4	US-09-862-154-1
11	340	25.1	509	3	US-09-039-555B-17
12	340	25.1	509	4	US-08-426-509A-18
13	340	25.1	509	4	US-09-457-040B-8
14	340	25.1	509	4	US-08-232-545-18
15	340	25.1	509	5	PCT-US95-05008-18
16	315.5	23.3	537	4	US-08-426-509A-11
17	315.5	23.3	537	4	US-08-232-545-11
18	315.5	23.3	537	5	PCT-US95-05008-11
19	315.5	23.3	543	4	US-08-426-509A-14
20	315.5	23.3	543	4	US-08-232-545-14
21	315.5	23.3	543	4	US-09-470-881-8
22	315.5	23.3	543	5	PCT-US95-05008-14
23	313.5	23.2	496	2	US-09-006-675-2
24	313.5	23.2	496	3	US-09-228-603A-2
25	312.5	23.1	529	4	US-08-426-509A-15
26	312.5	23.1	529	4	US-08-232-545-15
27	312.5	23.1	529	5	PCT-US95-05008-15

28	301	22.2	536	4	US-08-426-509A-12	Sequence 12, Appl
29	301	22.2	536	4	US-08-232-545-12	Sequence 12, Appl
30	301	22.2	536	5	PCT-US95-05008-12	Sequence 12, Appl
31	286.5	21.2	533	1	US-07-820-011A-2	Sequence 2, Appl
32	286.5	21.2	533	4	US-09-470-881-3	Sequence 3, Appl
33	286.5	21.2	533	5	PCT-US93-00445-2	Sequence 2, Appl
34	283.5	21.0	532	1	US-08-594-447-1	Sequence 1, Appl
35	283.5	21.0	532	2	US-08-665-647-1	Sequence 1, Appl
36	278	20.5	98	2	US-08-479-078-7	Sequence 7, Appl
37	276.5	20.4	536	1	US-07-820-011A-4	Sequence 4, Appl
38	276.5	20.4	536	4	US-08-426-509A-13	Sequence 13, Appl
39	276.5	20.4	536	4	US-08-232-545-13	Sequence 13, Appl
40	276.5	20.4	536	5	PCT-US93-00445-4	Sequence 4, Appl
41	276.5	20.4	536	5	PCT-US95-05008-13	Sequence 13, Appl
42	272	20.1	98	2	US-08-479-078-6	Sequence 6, Appl
43	266	19.7	108	5	PCT-US94-01840-6	Sequence 6, Appl
44	264	19.5	101	2	US-08-574-959A-5	Sequence 5, Appl
45	264	19.5	101	3	US-09-357-014-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1  
US-08-426-509A-16  
; Sequence 16, Application US/08426509A  
; Patent No. 6326469  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Gishizsky, Mikhail  
; APPLICANT: Sures, Irman G.  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
; TITLE OF INVENTION: TYROSINE KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York,  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/426,509A  
; FILING DATE: 21-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/232,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-0074-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 512 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: No. 6326469e  
US-08-426-509A-16

Query Match 27.4%; Score 370.5; DB 4; Length 512;  
Best Local Similarity 39.8%; Pred. No. 1.7e-31;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY 6 SRRKSLPSPSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPLTIYSED 65  
 DB 38 SNKQRPVE-SQLPGQRFQTKDPEEQGDIVVALYPYDGIHPDDLSPFKGKMKVLEEH 96  
 QY 66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLVEGLSREKABELLLPQNGAGFLIR 121  
 DB 97 GEWMAKSLTKKEGFIPSNYAKLNTLETEWFFKDIKDAERQQLAPGNSAGAFILR 156  
 QY 122 ESQTRGYSLSVRLSRPASWRIRHYRICHLDNGWLYISPRLTSPSLQALVDHYSELAD 181  
 DB 157 ESETLKGSSLSVRDPDPVHGDIKIRSLDNGGYISPRITPCISDMIKHYKQAD 216  
 QY 182 DICCLKEPCVLRAGPLPK 202  
 DB 217 GLCRRLEKACI-----SPKPK 233

RESULT 2  
 US-08-232-545-16  
 ; Sequence 16, Application US/08232545  
 ; Patent No. 6506578  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ulrich, Axel  
 ; APPLICANT: Gishitsky, Mikhail  
 ; APPLICANT: Sures, Iman G.  
 ; TITLE OF INVENTION: No. 6506578el Megakaryocytic Protein Tyrosine  
 ; TITLE OF INVENTION: Kinases  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232,545  
 ; FILING DATE: 22-APR-1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7683-050  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)790-9741  
 ; TELEFAX: (212)869-9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 16:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 512 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-08-232-545-16

Query Match 27.4%; Score 370.5; DB 4; Length 512;  
 Best Local Similarity 39.8%; Pred. No. 1.7e-31;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
 QY 6 SRRKSLPSPSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPLTIYSED 65  
 DB 38 SNKQRPVE-SQLPGQRFQTKDPEEQGDIVVALYPYDGIHPDDLSPFKGKMKVLEEH 96  
 QY 66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLVEGLSREKABELLLPQNGAGFLIR 121  
 DB 97 GEWMAKSLTKKEGFIPSNYAKLNTLETEWFFKDIKDAERQQLAPGNSAGAFILR 156

QY 122 ESQTRGYSLSVRLSRPASWRIRHYRICHLDNGWLYISPRLTSPSLQALVDHYSELAD 181  
 DB 157 ESETLKGSSLSVRDPDPVHGDIKIRSLDNGGYISPRITPCISDMIKHYKQAD 216  
 QY 182 DICCLKEPCVLRAGPLPK 202  
 DB 217 GLCRRLEKACI-----SPKPK 233

RESULT 3  
 PCT-US95-05008-16  
 ; Sequence 16, Application PC/TUS9505008  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sugen, Inc.  
 ; APPLICANT: 515 Galveston Drive  
 ; APPLICANT: Redwood City, California 94063-4720  
 ; APPLICANT: United States of America  
 ; APPLICANT: Wissenschaften E.V.  
 ; APPLICANT: Hofgarten Str. 2  
 ; APPLICANT: Munchen 80539  
 ; APPLICANT: Germany  
 ; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
 ; TITLE OF INVENTION: Kinases  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/05008  
 ; FILING DATE: 24-APR-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/232,545  
 ; FILING DATE: 22-APR-1994  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7683-074  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)790-9090  
 ; TELEFAX: (212)869-9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 16:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 512 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; PCT-US95-05008-16

Query Match 27.4%; Score 370.5; DB 5; Length 512;  
 Best Local Similarity 39.8%; Pred. No. 1.7e-31;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
 QY 6 SRRKSLPSPSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPLTIYSED 65  
 DB 38 SNKQRPVE-SQLPGQRFQTKDPEEQGDIVVALYPYDGIHPDDLSPFKGKMKVLEEH 96  
 QY 66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLVEGLSREKABELLLPQNGAGFLIR 121  
 DB 97 GEWMAKSLTKKEGFIPSNYAKLNTLETEWFFKDIKDAERQQLAPGNSAGAFILR 156

QY 122 ESQTRGYSLSVLRSPASWDRIHRYHICLDNGWLYISPRITFPLSQALVDHYSELAD 181  
Db 157 ESETLKGSPSLSVRDFPDVGHGDKIRSLDNGGYYISPRITFPCISDMIKHYKQAD 216  
QY 182 DICLLKEPCVLQAGPLGK 202  
Db 217 GLCRRLEKACI-----SPKPK 233  
RESULT 4  
US-08-426-509A-17  
; Sequence 17, Application US/08426509A  
; Patent No. 6326469  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Gishizky, Mikhail  
; APPLICANT: Sures, Iman G.  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
; TITLE OF INVENTION: TYROSINE KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York,  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/426.509A  
; FILING DATE: 21-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/232,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7883-0074-999  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 505 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
US-08-426-509A-17  
Query Match 26.6%; Score 360.5; DB 4; Length 505;  
Best Local Similarity 41.6%; Pred. No. 2e-30;  
Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
QY 12 PPSLSSSVQGGPVTMEAKRSKATAVALGSPAGPAELSLRLGEPPLTVISEDGDWTV 71  
Db 40 PGPNSHNS--NTPGIREAGSEDIIVVALDYEAITHHEDLSFQKGDQWVVLSESGEWKA 96  
QY 72 LSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKAEELLILPGNPGAFIRESQTRR 127  
Db 97 RSLATRKEGYIPSNVAVRVDLSLETBEWPFKISRKAERQLLAPGNMLGSPFMRDSEITK 156  
QY 128 GYSLSVLRSPASWDRIHRYHICLDNGWLYISPRITFPLSQALVDHYSELADICLL 187  
Db 157 GYSLSVRDYPDQGTGVKHYKIRTLDNGGFIISPRSTFSTLQELVDHYHKKNGDGLCQKL 216  
QY 188 KEPCV 192  
Db 217 SVPCM 221  
RESULT 6  
PCT-US95-05008-17  
; Sequence 17, Application PC/TUS9505008  
; GENERAL INFORMATION:

Db 217 SVPCM 221  
RESULT 5  
US-08-232-545-17  
; Sequence 17, Application US/08232545  
; Patent No. 6506578  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Gishizky, Mikhail  
; APPLICANT: Sures, Iman G.  
; TITLE OF INVENTION: No. 6506578el Megakaryocytic Protein Tyrosine  
; TITLE OF INVENTION: Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,545  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-050  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212)969-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 505 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-232-545-17  
Query Match 26.6%; Score 360.5; DB 4; Length 505;  
Best Local Similarity 41.6%; Pred. No. 2e-30;  
Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
QY 12 PPSLSSSVQGGPVTMEAKRSKATAVALGSPAGPAELSLRLGEPPLTVISEDGDWTV 71  
Db 40 PGPNSHNS--NTPGIREAGSEDIIVVALDYEAITHHEDLSFQKGDQWVVLSESGEWKA 96  
QY 72 LSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKAEELLILPGNPGAFIRESQTRR 127  
Db 97 RSLATRKEGYIPSNVAVRVDLSLETBEWPFKISRKAERQLLAPGNMLGSPFMRDSEITK 156  
QY 128 GYSLSVLRSPASWDRIHRYHICLDNGWLYISPRITFPLSQALVDHYSELADICLL 187  
Db 157 GYSLSVRDYPDQGTGVKHYKIRTLDNGGFIISPRSTFSTLQELVDHYHKKNGDGLCQKL 216  
QY 188 KEPCV 192  
Db 217 SVPCM 221  
RESULT 6  
PCT-US95-05008-17  
; Sequence 17, Application PC/TUS9505008  
; GENERAL INFORMATION:

APPLICANT: Sugen, Inc.  
 APPLICANT: 515 Galveston Drive  
 APPLICANT: Redwood City, California 94063-4720  
 APPLICANT: United States of America  
 APPLICANT: Wissenschaften E.V.  
 APPLICANT: Hofgarten Str. 2  
 APPLICANT: Munchen 80539  
 APPLICANT: Germany  
 TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
 TITLE OF INVENTION: Kinases  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/05008  
 FILING DATE: 24-APR-1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/232,545  
 FILING DATE: 22-APR-1994  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-074  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212)790-9090  
 TELEFAX: (212)869-9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 505 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 PCT-US95-05008-17

Query Match 26.6%; Score 360.5; DB 5; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30; Indels 7; Gaps 2;  
 Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 QY 12 PPSLSVSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGELPTIVSEDDWTV 71  
 Db 40 PCPNHNS--NTPGIREAGSEDIIVVAYDYEAIHEDLSFKQDQWVLEESGEWKA 96  
 QY 72 LSEVSGREYNIPSVHGVK----SHGWYEGLSREKAEELLPCNPGCAPLRESQTR 127  
 Db 97 RSLATRKEGYIFSNVAVRVDLSLETBEWFFKISRKAERQLLAPNMLGSMFIRDET 156  
 QY 128 GYSLSVLSRSPASWDRIHRIHCLDNGWLYISRLTFPSLQALVDHYSELADICLL 187  
 Db 157 GYSLSVLDYDPRQDVTGKYKIRTLDNGGFVSPRSTFTLQELVDHYKKGNDGLCQKL 216  
 QY 188 KEPCV 192  
 Db 217 SVPCM 221

RESULT 7  
 US-08-426-509A-19  
 ; Sequence 19, Application US/08426509A  
 ; Patent No. 6326469

GENERAL INFORMATION:  
 APPLICANT: Ullrich, Axel  
 APPLICANT: Gishizky, Mikhail  
 APPLICANT: Sures, Irman G.  
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
 TITLE OF INVENTION: TYROSINE KINASES  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York,  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/426,509A  
 FILING DATE: 21-APR-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/232,545  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-0074-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 499 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: No. 6326469e  
 US-08-426-509A-19  
 Query Match 25.5%; Score 344.5; DB 4; Length 499;  
 Best Local Similarity 36.0%; Pred. No. 1.1e-28;  
 Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;  
 QY 1 MGSLSFRRKSLPSLSVQGGPV-----TMEARSK 34  
 Db 1 MGLLSKRG-----VSEKKGWSPVKIRTDKAPPLPLVFNHAPPSPNQDPDEE 54  
 QY 35 ATAVALGSPAGGPAELSLRLGELPTIVSEDDWTVLSEVSGREYNIPSVHGVKVS--- 91  
 Db 55 RFVVALFDYAAVNDRLDQVKGKQLQVRSQDWMARSVLTGREGYVPSNFVAVETLE 114  
 QY 92 -HGWLVEGLSREKAEELLPCNPGCAPLRESQTRGYSLSVLSRSPASWDRIHRI 150  
 Db 115 VEKWFRTISRKAERQLLAPNMLGSMFIRDET 173  
 QY 151 HCLDNGWLYISRLTFPSLQALVDHYSELADICLLKEPCV 192  
 Db 174 RSLDNGGYISPRITFTLQALVQHYSKRGDGLCQKLTLPCV 215  
 RESULT 8  
 US-08-232-545-19  
 ; Sequence 19, Application US/08232545  
 ; Patent No. 6506578  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ullrich, Axel  
 ; APPLICANT: Gishizky, Mikhail  
 ; APPLICANT: Sures, Irman G.  
 ; TITLE OF INVENTION: No. 6506578e1 Megakaryocytic Protein Tyrosine

;; TITLE OF INVENTION: Kinases  
;; NUMBER OF SEQUENCES: 21  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Pennie & Edmonds  
;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: U.S.A.  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/232,545  
;; FILING DATE: 22-APR-1994  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Coruzzi, Laura A.  
;; REGISTRATION NUMBER: 30,742  
;; REFERENCE/DOCKET NUMBER: 7683-050  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)790-9090  
;; TELEFAX: (212)869-9741  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 499 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; PCT-US-08-232-545-19

Query Match 25.5%; Score 344.5; DB 4; Length 499;

Best Local Similarity 36.0%; Pred. No. 1.1e-28;  
Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;

QY 1 MGLSPRRKSLPSPSSSVQGGPV-----TWAEERSK 34  
DB 1 MGLSSKRG-----VSEKKGWSPVKIRTDKAPPLPLVFNHAPPSPNQPDDEE 54  
QY 35 ATAVALGSPAGPAELSLRLGEPLTVISDGDWTVLSEVSGREYNIPSVHVGKVS--- 91  
DB 55 RFVVALFDYAANDRDLDQVLKGEKQLVLRSTGDWMLARSLVTGREGVPSNFPVAPVETLE 114  
QY 92 -HGWLYEGLSREKAEELLPLPGNPGGAFLIRBSQTRRGYSLSVLSRPPASNDRIHYRI 150  
DB 115 VEKWFRTISRKDAERQQLAPNKGAGFLIRSESNKGAFSLSVK-DITTOGEVVVKHYKI 173  
QY 151 HCLDNGWLYISPRLTTPPSLQALVDHYSELADDCILLKEPCV 192  
DB 174 RSLDNGGYIISPRITFTLQALVQHYSKKGDGLCQKLTLCV 215

## RESULT 9

PCT-US95-05008-19  
; Sequence 19, Application PC/TUS9505008  
; GENERAL INFORMATION:  
; APPLICANT: Sugen, Inc.  
; APPLICANT: 515 Galveston Drive  
; APPLICANT: Redwood City, California 94063-4720  
; APPLICANT: United States of America  
; APPLICANT: Wissenschaften E.V.  
; APPLICANT: Hofgarten Str. 2  
; APPLICANT: Munchen 80539  
; APPLICANT: Germany  
; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds

;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: U.S.A.  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/05008  
;; FILING DATE: 24-APR-1995  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/232,545  
;; FILING DATE: 22-APR-1994  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Coruzzi, Laura A.  
;; REGISTRATION NUMBER: 30,742  
;; REFERENCE/DOCKET NUMBER: 7683-074  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)790-9090  
;; TELEFAX: (212)869-9741  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 499 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; PCT-US95-05008-19

Query Match 25.5%; Score 344.5; DB 5; Length 499;

Best Local Similarity 36.0%; Pred. No. 1.1e-28;  
Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;

QY 1 MGLSPRRKSLPSPSSSVQGGPV-----TWAEERSK 34  
DB 1 MGLSSKRG-----VSEKKGWSPVKIRTDKAPPLPLVFNHAPPSPNQPDDEE 54  
QY 35 ATAVALGSPAGPAELSLRLGEPLTVISDGDWTVLSEVSGREYNIPSVHVGKVS--- 91  
DB 55 RFVVALFDYAANDRDLDQVLKGEKQLVLRSTGDWMLARSLVTGREGVPSNFPVAPVETLE 114  
QY 92 -HGWLYEGLSREKAEELLPLPGNPGGAFLIRBSQTRRGYSLSVLSRPPASNDRIHYRI 150  
DB 115 VEKWFRTISRKDAERQQLAPNKGAGFLIRSESNKGAFSLSVK-DITTOGEVVVKHYKI 173  
QY 151 HCLDNGWLYISPRLTTPPSLQALVDHYSELADDCILLKEPCV 192  
DB 174 RSLDNGGYIISPRITFTLQALVQHYSKKGDGLCQKLTLCV 215

## RESULT 10

US-09-862-154-1  
; Sequence 1, Application US/09862154  
; Patent No. 6589758  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Xiaolian  
; TITLE OF INVENTION: Crystal of a Kinase-ligand Complex and Methods of Use  
; FILE REFERENCE: Atty. Docket No. 6589758: A-749  
; CURRENT APPLICATION NUMBER: US/09/862,154  
; CURRENT FILING DATE: 2001-05-21  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 508  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-862-154-1

Query Match 25.1%; Score 340; DB 4; Length 508;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
QY 26 VTMEARSKAT-----AVALGSPFAGGPAELSLRLGEPPLTVSDGDGDMWTVLSEVSGRE 79  
DB 48 VTYEGSNPPASPLQDNLVIALHSYEPHSDGDLGFEKGEOQLRILEQSGEWWKQAQSLTTQGE 107  
QY 80 YNIPSVHVGVKVS---HGMWLYEGLSREKAEELLLPQNPFGAFLIRESOTRSGYSLSVR 135  
DB 108 GFIPFNFVAKANLEPEPWFKNLSRDAERQQLAPGNTHGSLFIRESESTAGSFSLSVR 167  
QY 136 LSRPASWDRIRHYRIHCLDNGWLYISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 195  
DB 168 DFDQNGQEVVVKYKIRNLNDNGGFIISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 227

RESULT 11  
US-09-039-555B-17  
; Sequence 17, Application US/09039555B  
; Patent No. 6033856  
; GENERAL INFORMATION:  
; APPLICANT: Koerner, Kathrin  
; APPLICANT: Mueller, Rolf  
; APPLICANT: Sadlaczek, Hans-Harald  
; TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS  
; PREPARATION AND USE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,555B  
; FILING DATE: 16-MAR-1998  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 19710643.9  
; FILING DATE: 14-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 016779/0131  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 509 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: linear  
; MOLECULE TYPE: protein  
US-09-039-555B-17

Query Match 25.1%; Score 340; DB 3; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
QY 26 VTMEARSKAT-----AVALGSPFAGGPAELSLRLGEPPLTVSDGDGDMWTVLSEVSGRE 79  
DB 49 VTYEGSNPPASPLQDNLVIALHSYEPHSDGDLGFEKGEOQLRILEQSGEWWKQAQSLTTQGE 108

QY 80 YNIPSVHVGVKVS---HGMWLYEGLSREKAEELLLPQNPFGAFLIRESOTRSGYSLSVR 135  
DB 109 GFIPFNFVAKANLEPEPWFKNLSRDAERQQLAPGNTHGSLFIRESESTAGSFSLSVR 168  
QY 136 LSRPASWDRIRHYRIHCLDNGWLYISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 195  
DB 169 DFDQNGQEVVVKYKIRNLNDNGGFIISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 228

RESULT 12  
US-08-426-509A-18  
; Sequence 18, Application US/08426509A  
; Patent No. 6326469  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Gishizky, Mikhail  
; APPLICANT: Suresh, Iman G.  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
; TYROSINE KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York,  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/426,509A  
; FILING DATE: 21-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/232,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-0074-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-869-9741  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 509 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
US-08-426-509A-18

Query Match 25.1%; Score 340; DB 4; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

QY 26 VTMEARSKAT-----AVALGSPFAGGPAELSLRLGEPPLTVSDGDGDMWTVLSEVSGRE 79  
DB 49 VTYEGSNPPASPLQDNLVIALHSYEPHSDGDLGFEKGEOQLRILEQSGEWWKQAQSLTTQGE 108  
QY 80 YNIPSVHVGVKVS---HGMWLYEGLSREKAEELLLPQNPFGAFLIRESOTRSGYSLSVR 135  
DB 109 GFIPFNFVAKANLEPEPWFKNLSRDAERQQLAPGNTHGSLFIRESESTAGSFSLSVR 168  
QY 136 LSRPASWDRIRHYRIHCLDNGWLYISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 195  
DB 169 DFDQNGQEVVVKYKIRNLNDNGGFIISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 228

RESULT 13



US-09-457-040B-8  
; Sequence 8, Application US/09457040B  
; Patent No. 6387641  
; GENERAL INFORMATION:  
; APPLICANT: Vertex Pharmaceuticals Incorporated  
; APPLICANT: Bellon, Steve  
; TITLE OF INVENTION: Crystallized P38 Complexes  
; FILE REFERENCE: VPI/98-14  
; CURRENT APPLICATION NUMBER: US/09/457,040B  
; CURRENT FILING DATE: 1999-12-08  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Human  
US-09-457-040B-8

Query Match 25.1%; Score 340; DB 4; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
  
QY 26 VTMEARSKAT-----AVALGSPAGGPAELSLRLGPEPLTIVSEDDGWMTVLSEVSGRE 79  
Db 49 VTVEGSPASPQDNLVIALHSYEPHSDGLGFEKGEQLRILEQSGEWKQAQSLTTQGE 108  
  
QY 80 YNIPSVHVGVKVS-----HGMWYEGLSREKAEELLLPGNPGGAFILRESQTRRGSYSLSVR 135  
Db 109 GTIFFNFVAKANSLPEPEPFKKNLSKDAERQLLAPGNTHGSLFIRESESTAGSFLSVR 169  
  
QY 136 LSRPASWDRIHYRIHCLDNGWLYISPRITFFPSLQALVDHYSELADDDICLLKEPCVLQR 195  
Db 169 DFDQNGQEVVKYKIRNLNDGGFYISPRITFFGLHELVRHYTNASDGLCTRLSRPCQOK 228

RESULT 14  
US-08-232-545-18  
; Sequence 18, Application US/08232545  
; Patent No. 6506578  
; GENERAL INFORMATION:  
; APPLICANT: Ulrich, Axel  
; APPLICANT: Gishizky, Mikhail  
; APPLICANT: Sures, Irmann G.  
; TITLE OF INVENTION: Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,545  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-050  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212)869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 509 amino acids

; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-232-545-18  
  
Query Match 25.1%; Score 340; DB 4; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
  
QY 26 VTMEARSKAT-----AVALGSPAGGPAELSLRLGPEPLTIVSEDDGWMTVLSEVSGRE 79  
Db 49 VTVEGSPASPQDNLVIALHSYEPHSDGLGFEKGEQLRILEQSGEWKQAQSLTTQGE 108  
  
QY 80 YNIPSVHVGVKVS-----HGMWYEGLSREKAEELLLPGNPGGAFILRESQTRRGSYSLSVR 135  
Db 109 GTIFFNFVAKANSLPEPEPFKKNLSKDAERQLLAPGNTHGSLFIRESESTAGSFLSVR 168  
  
QY 136 LSRPASWDRIHYRIHCLDNGWLYISPRITFFPSLQALVDHYSELADDDICLLKEPCVLQR 195  
Db 169 DFDQNGQEVVKYKIRNLNDGGFYISPRITFFGLHELVRHYTNASDGLCTRLSRPCQOK 228

RESULT 15  
PCT-US95-05008-18  
; Sequence 18, Application PC/TUS9505008  
; GENERAL INFORMATION:  
; APPLICANT: Sugen, Inc.  
; APPLICANT: 515 Galveston Drive  
; APPLICANT: Redwood City, California 94063-4720  
; APPLICANT: United States of America  
; APPLICANT: Wissenschaften E.V.  
; APPLICANT: Hofgarten Str. 2  
; APPLICANT: Munchen 80539  
; APPLICANT: Germany  
; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
; TITLE OF INVENTION: Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05008  
; FILING DATE: 24-APR-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/232,545  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212)869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 509 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein

PCT-US95-05008-18

```

Query Match      25.1%; Score 340; DB 5; Length 509;
Best Local Similarity 40.6%; Pred. No. 3.5e-28;
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

Qy 26 VTWEAERSKAT-----AVALGSPAGGPAELSLRLGEPITIVSEDDGDMWTVLSEVSGRE 79
Db 49 VTVEGSPNPASPLQDNLVIALHSYEPESHGDLGFEKGEQLRILEQSGEWWKQSLTTGQE 108
Qy 80 YNIPSVHVGVKS----HGWLVEGLSREKAEELLLEFNPQGAFLIRESQTRRGYSLSYR 135
Db 109 GFIPFNFAKANSLPEPEPFKNLSRKDAERQLLAPGNTHGSLFIRESESTAGSFSLSVR 168
Qy 136 LSRPASWDRIHYRIHCLDNGWLYISPRITFPSLQALVDHYSELADDICLLKEPCVLQR 195
Db 169 DFDQNGEVVVKHYKIRNLNNGGFYISPRITFPGLHELVRHYTNASDGLCTRLSRPCQTK 228

```

Search completed: February 20, 2004, 09:34:03  
Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:32:17 ; Search time 593 Seconds  
(without alignments)  
92.936 Million cell updates/sec

Title: US-09-939-853A-75  
Perfect score: 1353  
Sequence: 1 MGSILPSRKSLPSPSLSSV.....RESLSFVSLNDEAVSLDDA 261

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09C\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10C\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1347	99.6	261	14	US-10-043-649-2
2	826	61.0	159	9	US-09-867-550-954
3	586	43.3	113	9	US-09-867-550-1916
4	481.5	35.6	276	9	US-09-870-759-64
5	481.5	35.6	276	10	US-09-751-708A-64
6	481.5	35.6	276	14	US-10-043-649-3
7	452.5	33.4	96	9	US-09-867-550-952
8	370.5	27.4	511	15	US-10-394-322A-42
9	370.5	27.4	512	9	US-09-977-269-16
10	370.5	27.4	512	9	US-09-977-269-16
11	370.5	27.4	512	15	US-10-116-275-16
12	370.5	27.4	512	15	US-10-116-275-16
13	360.5	26.6	505	9	US-09-977-269-17
14	360.5	26.6	505	9	US-09-977-260-17
15	360.5	26.6	505	10	US-09-977-261-17

16	360.5	26.6	505	15	US-10-193-720-2	Sequence 2, Appli
17	360.5	26.6	526	15	US-10-394-322A-31	Sequence 31, Appli
18	352.5	26.1	504	15	US-10-394-322A-4	Sequence 4, Appli
19	352.5	26.1	505	9	US-09-771-161A-186	Sequence 186, Appl
20	348	25.7	509	15	US-10-366-288-28	Sequence 28, Appl
21	344.5	25.5	499	9	US-09-977-269-19	Sequence 19, Appli
22	344.5	25.5	499	9	US-09-977-260-19	Sequence 19, Appli
23	344.5	25.5	499	10	US-09-977-261-19	Sequence 19, Appli
24	340	25.1	508	15	US-10-394-322A-41	Sequence 41, Appli
25	340	25.1	509	9	US-09-977-269-18	Sequence 18, Appli
26	340	25.1	509	9	US-09-977-260-18	Sequence 18, Appli
27	340	25.1	509	10	US-09-977-261-18	Sequence 18, Appli
28	340	25.1	509	14	US-10-212-346-1	Sequence 1, Appli
29	322.5	23.8	454	9	US-09-771-161A-95	Sequence 95, Appli
30	318.5	23.5	537	9	US-09-771-161A-212	Sequence 212, App
31	318.5	23.5	537	9	US-09-771-161A-213	Sequence 213, App
32	315.5	23.3	311	9	US-09-771-161A-121	Sequence 121, App
33	315.5	23.3	387	9	US-09-771-161A-122	Sequence 122, App
34	315.5	23.3	536	15	US-10-394-322A-30	Sequence 30, Appl
35	315.5	23.3	537	9	US-09-977-269-11	Sequence 11, Appli
36	315.5	23.3	537	9	US-09-977-260-11	Sequence 11, Appli
37	315.5	23.3	537	10	US-09-977-261-11	Sequence 11, Appli
38	315.5	23.3	543	9	US-09-977-269-14	Sequence 14, Appli
39	315.5	23.3	543	9	US-09-977-260-14	Sequence 14, Appli
40	315.5	23.3	543	10	US-09-977-261-14	Sequence 14, Appli
41	315.5	23.3	543	14	US-10-298-377A-4	Sequence 4, Appli
42	315.5	23.3	543	15	US-10-394-322A-68	Sequence 68, Appli
43	312.5	23.1	529	9	US-09-977-269-15	Sequence 15, Appli
44	312.5	23.1	529	9	US-09-977-260-15	Sequence 15, Appli
45	312.5	23.1	529	10	US-09-977-261-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1

US-10-043-649-2  
; Sequence 2, Application US/10043649  
; Publication No. US20030059924A1  
; GENERAL INFORMATION:  
; APPLICANT: Holland, Sacha J.  
; APPLICANT: Mendenhall, Marcy K.  
; APPLICANT: Pardo, Jorge  
; APPLICANT: Spencer, Collin  
; APPLICANT: Fu, C. Alan  
; APPLICANT: Luo, Ying  
; APPLICANT: Payan, Donald G.  
; APPLICANT: Mancebo, Helena S.Y.  
; APPLICANT: Wu, Jun  
; APPLICANT: Zhou, Xiulan  
; APPLICANT: Shen, Mary  
; APPLICANT: Liao, X. Charlene  
; APPLICANT: Sheng, Ning  
; TITLE OF INVENTION: Cloning of a No. US20030059924A1e1 Inhibitor of Antigen-receptor ;  
; TITLE OF INVENTION: Retroviral-based Functional Screen  
; FILE REFERENCE: A-70219-1/RMS/DHR  
; CURRENT APPLICATION NUMBER: US/10/043,649  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/260,953  
; PRIOR FILING DATE: 2001-01-10  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-043-649-2

Query Match 99.6%; Score 1347; DB 14; Length 261;  
Best Local Similarity 99.6%; Pred. No. 2e-124;  
Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSILPSRKSLPSPSLSSVQQGGPTMEARSKATAVALGSPAGPAELSLRLCEPLT 60

Db 1 MGSLPSRRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPFLT 60  
Qy 61 IVSEDDGNTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLLPGNPGGAFLI 120  
Db 61 IVSEDDGNTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLLPGNPGGAFLI 120  
Qy 121 RESQTRRGYSLSVRLSRPASWDRIHRHRIHCLDNGWLYISPRLTSPSLQALVDHYSELA 180  
Db 121 RESQTRRGYSLSVRLSRPASWDRIHRHRIHCLDNGWLYISPRLTSPSLQALVDHYSELA 180  
Qy 181 DDICCLLKPCVLRAGPLPGKDIPLPVTVQRTPLANKKELDSLLPSEATGEESLLSG 240  
Db 181 DDICCLLKPCVLRAGPLPGKDIPLPVTVQRTPLANKKELDSLLPSEATGEESLLSG 240  
Qy 241 LRESLSFYISLNDSEAVSLDDA 261  
Db 241 LRESLSFYISLNDSEAVSLDDA 261

RESULT 2  
US-09-867-550-954  
; Sequence 954, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fud, Pamela  
; APPLICANT: Conley, Debbie  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 954  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-867-550-954

Query Match 61.0%; Score 826; DB 9; Length 159;  
Best Local Similarity 99.4%; Pred. No. 2.3e-73;  
Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MGSLPSRRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPFLT 60  
Db 1 MGSLPSRRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPFLT 60  
Qy 61 IVSEDDGNTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLLPGNPGGAFLI 120  
Db 61 IVSEDDGNTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLLPGNPGGAFLI 120  
Qy 121 RESQTRRGYSLSVRLSRPASWDRIHRHRIHCLDNGWLY 159  
Db 121 RESQTRRGYSLSVRLSRPASWDRIHRHRIHCLDNGWLY 159

RESULT 3  
US-09-867-550-1916  
; Sequence 1916, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fud, Pamela  
; APPLICANT: Conley, Debbie  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and

; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1916  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: VARIANT  
; LOCATION: (1)  
; OTHER INFORMATION: Wherein Xaa may be any one of Arg or Gly or Trp  
US-09-867-550-1916

Query Match 43.3%; Score 586; DB 9; Length 113;  
Best Local Similarity 100.0%; Pred. No. 6.8e-50;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 150 IHCLDNGWLYISPRLTSPSLQALVDHYSELAADDICCLLKPCVLRAGPLPGKDIPLPVT 209  
Db 2 IHCLDNGWLYISPRLTSPSLQALVDHYSELAADDICCLLKPCVLRAGPLPGKDIPLPVT 61  
Qy 210 VQRTPLANKKELDSLLPSEATGEESLLSGLRESLSFYISLNDSEAVSLDDA 261  
Db 62 VQRTPLANKKELDSLLPSEATGEESLLSGLRESLSFYISLNDSEAVSLDDA 113

RESULT 4  
US-09-870-759-64  
; Sequence 64, Application US/09870759  
; Patent No. US20020177551A1  
; GENERAL INFORMATION:  
; APPLICANT: Ferman, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; CURRENT FILING DATE: 2002-01-14  
; PRIOR APPLICATION NUMBER: US 60/208,128  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 64  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-759-64

Query Match 35.6%; Score 481.5; DB 9; Length 276;  
Best Local Similarity 40.3%; Pred. No. 4.8e-39;  
Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;  
Qy 9 KSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPFLTIVSEDDGW 68  
Db 6 KSTPAPA-----ERPLNPEGLSDSFLAVLSDFSPDISPPIFRGKLRVISEGGW 58  
Qy 69 WTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLLPGNPGGAFLIREQTRRG 128  
Db 59 WKALISLTGRESYIPGICVARVTHGWLFEGLORDKAEELLQLPDTKVGSMIRESETKKG 118  
Qy 129 SYSLSVRLSRPASWDRIHRHRIHCLDNGWLYISPRLTSPSLQALVDHYSELAADDICCLLK 188  
Db 119 FYSLSVR-----HRQVKHYRIFRLPNWYIISPRLTFCQLEDLVNHYSEVADGJCCVLT 172  
Qy 189 EPCVLRAGPLPGKDIPLPVTVQRTPLANKKELDSLLPSEATG-----EESLISEGL 241  
Db 173 TPCLTOSTAAPAVRASSSPVTLRQKTVDMRWVR---LQEDPEGTENPLGVDESLSYGL 229  
Qy 242 RESLSFYISLNDSE 254

Db 230 RESIASYLSLTSE 242

RESULT 5

US-09-751-708A-64

; Sequence 64, Application US/09751708A

; Publication No. US20030157113A1

; GENERAL INFORMATION:

; APPLICANT: TERMAN, David S

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE

; FILE REFERENCE: 751708

; CURRENT FILING DATE: 2002-10-15

; PRIOR APPLICATION NUMBER: US 60/173,371

; PRIOR FILING DATE: 1999-12-28

; NUMBER OF SEQ ID NOS: 166

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 64

; LENGTH: 276

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-751-708A-64

Query Match 35.6%; Score 481.5; DB 10; Length 276;

Best Local Similarity 40.3%; Pred. No. 4.8e-39;

Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;

QY 9 KSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPITIVSEDDGW 68

Db 6 KSTPAPA-----ERPLNPEGLSDDFLAVLSQVPSDPISPPFRGEKLRVSDGGW 58

QY 69 WTVLSEVSGREYNIPSHVGVKSHGWLVEGLSREKAEELLPLGNPGGAFLIRESQTRG 128

Db 59 WKALSISTGRESYIPGICVARVTHGWLFEGLGRDKAEELLQPDPTKVGSPFMIRESETKKG 118

QY 129 SYSLSVRLSRPASWDRIRHYRTHCLDNGWLYTSPLTTPSLQALVDHYSSELADDCCLLK 188

Db 119 FYLSLVR-----HRQVKHYRIFRLPNWYIISPLTTCQLEDLVNHYSEVADGLCCVLT 172

QY 189 EPCVLQAGPLPKDIPVTVTQRTPLNWKELDSSLFSEAAATG-----EESLLSEGL 241

Db 173 TPCLTQSTAAPAVRASSSPVTLRQKTVDWRRVSR---LQEDPEGTENPLGVDESLSFYGL 229

QY 242 RESLSFYISLNDE 254

Db 230 RESIASYLSLTSE 242

RESULT 6

US-10-043-649-3

; Sequence 3, Application US/10043649

; Publication No. US20030059924A1

; GENERAL INFORMATION:

; APPLICANT: Holland, Sacha J.

; APPLICANT: Mendenhall, Marcy K.

; APPLICANT: Pardo, Jorge

; APPLICANT: Spencer, Collin

; APPLICANT: Fu, C. Alan

; APPLICANT: Luo, Ying

; APPLICANT: Payan, Donald G.

; APPLICANT: Mancebo, Helena S.Y.

; APPLICANT: Wu, Jun

; APPLICANT: Zhou, Xiulan

; APPLICANT: Shen, Mary

; APPLICANT: Liao, X. Charlene

; APPLICANT: Sheng, Ning

; TITLE OF INVENTION: Cloning of a No. US20030059924A1a1 Inhibitor of Antigen-receptor

; FILE REFERENCE: Retroviral-based Functional Screen

; FILE REFERENCE: A-70219-1/RMS/DHR

; CURRENT FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 60/260,953

; PRIOR FILING DATE: 2001-01-10

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 3

; LENGTH: 276

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-043-649-3

Query Match 35.6%; Score 481.5; DB 14; Length 276;

Best Local Similarity 40.3%; Pred. No. 4.8e-39;

Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;

QY 9 KSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPITIVSEDDGW 58

Db 6 KSTPAPA-----ERPLNPEGLSDDFLAVLSQVPSDPISPPFRGEKLRVSDGGW 58

QY 69 WTVLSEVSGREYNIPSHVGVKSHGWLVEGLSREKAEELLPLGNPGGAFLIRESQTRG 128

Db 59 WKALSISTGRESYIPGICVARVTHGWLFEGLGRDKAEELLQPDPTKVGSPFMIRESETKKG 118

QY 129 SYSLSVRLSRPASWDRIRHYRTHCLDNGWLYTSPLTTPSLQALVDHYSSELADDCCLLK 188

Db 119 FYLSLVR-----HRQVKHYRIFRLPNWYIISPLTTCQLEDLVNHYSEVADGLCCVLT 172

QY 189 EPCVLQAGPLPKDIPVTVTQRTPLNWKELDSSLFSEAAATG-----EESLLSEGL 241

Db 173 TPCLTQSTAAPAVRASSSPVTLRQKTVDWRRVSR---LQEDPEGTENPLGVDESLSFYGL 229

QY 242 RESLSFYISLNDE 254

Db 230 RESIASYLSLTSE 242

RESULT 7

US-09-867-550-952

; Sequence 952, Application US/09867550

; Patent No. US20020082206A1

; GENERAL INFORMATION:

; APPLICANT: Leach, Martin D.

; APPLICANT: Menhaban, Fuad,

; APPLICANT: Conley, Pamela

; APPLICANT: Law, Debbie

; APPLICANT: Topper, James

; TITLE OF INVENTION: No. US20020082206A1a1 Polynucleotides from Atherogenic Cells and i

; FILE REFERENCE: 21402-013 (Cura-313)

; CURRENT FILING DATE: 2001-09-20

; PRIOR APPLICATION NUMBER: USN 60/208,427

; PRIOR FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 2125

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 952

; LENGTH: 96

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-867-550-952

Query Match 33.4%; Score 452.5; DB 9; Length 96;

Best Local Similarity 76.8%; Pred. No. 7.9e-37;

Matches 96; Conservative 0; Mismatches 0; Indels 29; Gaps 1;

QY 1 MGSLSRKRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPIT 60

Db 1 MGSLSRKRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPIT 60

QY 61 IVSEDDGWTVLSEVSGREYNIPSHVGVKSHGWLVEGLSREKAEELLPLGNPGGAFLI 120

Db 61 IVSE-----WLVEGLSREKAEELLPLGNPGGAFLI 91

QY 121 RESQT 125

Db 92 RESQT 96

QY 66 GDWWTVLSEVSGREYNIPSVHGVK-----SHGWLVEGLSREKAEELLILLPGNPGGAFLIR 121  
 Db 97 GEWWAKSLITKKEGFIPSNVAKLNTLEETEWFKKDITRKDAERQLLAFGNSAGAFILR 156  
 QY 122 ESQTRRGYSLSVRLSRPASWDRIHRYHICLDNGWLYISPRITFPLSLQALVDHYSELAD 181  
 Db 157 ESETLKGFSLSVRDFFPVHGDVIRKIRSLDNGGYYISPRITFPCISDMIKHYKQAD 216  
 QY 182 DICLLKEPCVLQAGPLPGK 202  
 Db 217 GLCRRLEKACI-----SPKPK 233

RESULT 10

US-09-977-260-16  
 ; Sequence 16, Application US/09977260  
 ; Publication No. US20020192790A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ULLRICH, AXEL  
 ; APPLICANT: GISHIZKY, MIKHAIL  
 ; APPLICANT: SURES, IRMINGARD  
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
 ; FILE REFERENCE: 038602/1260  
 ; CURRENT APPLICATION NUMBER: US/09/977,260  
 ; PRIOR FILING DATE: 2001-10-16  
 ; PRIOR APPLICATION NUMBER: 09/232,545  
 ; PRIOR FILING DATE: 1994-04-22  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 16  
 ; LENGTH: 512  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-977-260-16

Query Match 27.4%; Score 370.5; DB 9; Length 512;  
 Best Local Similarity 39.8%; Pred. No. 1e-27;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY 6 SRRKSLPSPSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGELPTIVSED 65  
 Db 38 SNKQRPVPE-SQLLPQRFQTKDPEEQDIVVALYPYDGIHPDDLSFKKGEKMKVLEH 96  
 QY 66 GDWWTVLSEVSGREYNIPSVHGVK-----SHGWLVEGLSREKAEELLILLPGNPGGAFLIR 121  
 Db 97 GEWWAKSLITKKEGFIPSNVAKLNTLEETEWFKKDITRKDAERQLLAFGNSAGAFILR 156  
 QY 122 ESQTRRGYSLSVRLSRPASWDRIHRYHICLDNGWLYISPRITFPLSLQALVDHYSELAD 181  
 Db 157 ESETLKGFSLSVRDFFPVHGDVIRKIRSLDNGGYYISPRITFPCISDMIKHYKQAD 216  
 QY 182 DICLLKEPCVLQAGPLPGK 202  
 Db 217 GLCRRLEKACI-----SPKPK 233

RESULT 11

US-09-977-261-16  
 ; Sequence 16, Application US/09977261  
 ; Publication No. US20030054527A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ULLRICH, AXEL  
 ; APPLICANT: GISHIZKY, MIKHAIL  
 ; APPLICANT: SURES, IRMINGARD  
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
 ; FILE REFERENCE: 038602/1259  
 ; CURRENT APPLICATION NUMBER: US/09/977,261  
 ; PRIOR FILING DATE: 2001-10-16  
 ; PRIOR APPLICATION NUMBER: 08/232,545  
 ; PRIOR FILING DATE: 1994-04-22  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn Ver. 2.1

QY 66 GDWWTVLSEVSGREYNIPSVHGVK-----SHGWLVEGLSREKAEELLILLPGNPGGAFLIR 121  
 Db 97 GEWWAKSLITKKEGFIPSNVAKLNTLEETEWFKKDITRKDAERQLLAFGNSAGAFILR 156  
 QY 122 ESQTRRGYSLSVRLSRPASWDRIHRYHICLDNGWLYISPRITFPLSLQALVDHYSELAD 181  
 Db 157 ESETLKGFSLSVRDFFPVHGDVIRKIRSLDNGGYYISPRITFPCISDMIKHYKQAD 216  
 QY 182 DICLLKEPCVLQAGPLPGK 202  
 Db 217 GLCRRLEKACI-----SPKPK 233

RESULT 8

US-10-394-322A-42  
 ; Sequence 42, Application US/10394322A  
 ; Publication No. US20030232391A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SUNESIS PHARMACEUTICALS, INC.  
 ; APPLICANT: Prescott, John C.  
 ; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
 ; FILE REFERENCE: 39750-0006 US  
 ; CURRENT APPLICATION NUMBER: US/10/394,322A  
 ; CURRENT FILING DATE: 2003-03-20  
 ; PRIOR APPLICATION NUMBER: US 60/366,892  
 ; PRIOR FILING DATE: 2002-03-21  
 ; NUMBER OF SEQ ID NOS: 70  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 42  
 ; LENGTH: 511  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-394-322A-42

Query Match 27.4%; Score 370.5; DB 15; Length 511;  
 Best Local Similarity 39.8%; Pred. No. 1e-27;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY 6 SRRKSLPSPSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGELPTIVSED 65  
 Db 37 SNKQRPVPE-SQLLPQRFQTKDPEEQDIVVALYPYDGIHPDDLSFKKGEKMKVLEH 95  
 QY 66 GDWWTVLSEVSGREYNIPSVHGVK-----SHGWLVEGLSREKAEELLILLPGNPGGAFLIR 121  
 Db 96 GEWWAKSLITKKEGFIPSNVAKLNTLEETEWFKKDITRKDAERQLLAFGNSAGAFILR 155  
 QY 122 ESQTRRGYSLSVRLSRPASWDRIHRYHICLDNGWLYISPRITFPLSLQALVDHYSELAD 181  
 Db 156 ESETLKGFSLSVRDFFPVHGDVIRKIRSLDNGGYYISPRITFPCISDMIKHYKQAD 215  
 QY 182 DICLLKEPCVLQAGPLPGK 202  
 Db 216 GLCRRLEKACI-----SPKPK 232

RESULT 9

US-09-977-269-16  
 ; Sequence 16, Application US/09977269  
 ; Patent No. US2002082037A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ULLRICH, AXEL  
 ; APPLICANT: GISHIZKY, MIKHAIL  
 ; APPLICANT: SURES, IRMINGARD  
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
 ; FILE REFERENCE: 038602/1260  
 ; CURRENT APPLICATION NUMBER: US/09/977,269  
 ; PRIOR FILING DATE: 2001-10-16  
 ; PRIOR APPLICATION NUMBER: 08/232,545  
 ; PRIOR FILING DATE: 1994-04-22  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 16  
 ; LENGTH: 512  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-977-269-16

Query Match 27.4%; Score 370.5; DB 9; Length 512;  
 Best Local Similarity 39.8%; Pred. No. 1e-27;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY 6 SRRKSLPSPSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGELPTIVSED 65  
 Db 38 SNKQRPVPE-SQLLPQRFQTKDPEEQDIVVALYPYDGIHPDDLSFKKGEKMKVLEH 96

```
; SEQ ID NO 16
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-261-16

Query Match      27.4%; Score 370.5; DB 10; Length 512;
Best Local Similarity 39.8%; Pred. No. 1e-27;
Matches      80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY  6 SRKSLPSPSSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSD 65
DB  38 SNKQRPVPE-SQLPGQRFQTKPEEGDIIIVALYPDGHDPDLSPFKGKMKVLEH 96
QY  66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIR 121
DB  97 GEMWAKSLITKEGFIPIVAKLNTLEETEEWFKDITRKDAERQLLAPGNSAGAFLIR 156
QY  122 ESQTRGYSLSVRLSRPASWDRIRHYRICHLDNGWLYISPRITFPPSLQALVDHYSELAD 181
DB  157 ESETLKGFSLSVRDFFDVHGDIVKHYSRLDNGGYISPRITFPICISDMIKHYKQKQAD 216
QY  182 DICLLKEPCVLQAGPLPGK 202
DB  217 GLCRLEKACI-----SPKPK 233

RESULT 12
US-10-116-275-162
; Sequence 162, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 162
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-162

Query Match      27.4%; Score 370.5; DB 15; Length 512;
Best Local Similarity 39.8%; Pred. No. 1e-27;
Matches      80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY  6 SRKSLPSPSSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSD 65
DB  38 SNKQRPVPE-SQLPGQRFQTKPEEGDIIIVALYPDGHDPDLSPFKGKMKVLEH 96
QY  66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIR 121
DB  97 GEMWAKSLITKEGFIPIVAKLNTLEETEEWFKDITRKDAERQLLAPGNSAGAFLIR 156
QY  122 ESQTRGYSLSVRLSRPASWDRIRHYRICHLDNGWLYISPRITFPPSLQALVDHYSELAD 181
DB  157 ESETLKGFSLSVRDFFDVHGDIVKHYSRLDNGGYISPRITFPICISDMIKHYKQKQAD 216
QY  182 DICLLKEPCVLQAGPLPGK 202
DB  217 GLCRLEKACI-----SPKPK 233

RESULT 13
US-10-116-275-162
; Sequence 162, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 162
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-162

Query Match      27.4%; Score 370.5; DB 15; Length 512;
Best Local Similarity 39.8%; Pred. No. 1e-27;
Matches      80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY  6 SRKSLPSPSSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSD 65
DB  38 SNKQRPVPE-SQLPGQRFQTKPEEGDIIIVALYPDGHDPDLSPFKGKMKVLEH 96
QY  66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIR 121
DB  97 GEMWAKSLITKEGFIPIVAKLNTLEETEEWFKDITRKDAERQLLAPGNSAGAFLIR 156
QY  122 ESQTRGYSLSVRLSRPASWDRIRHYRICHLDNGWLYISPRITFPPSLQALVDHYSELAD 181
DB  157 ESETLKGFSLSVRDFFDVHGDIVKHYSRLDNGGYISPRITFPICISDMIKHYKQKQAD 216
QY  182 DICLLKEPCVLQAGPLPGK 202
DB  217 GLCRLEKACI-----SPKPK 233
```

```
US-09-977-269-17
; Sequence 17, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,269
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-269-17

Query Match      26.6%; Score 360.5; DB 9; Length 505;
Best Local Similarity 41.6%; Pred. No. 9.6e-27;
Matches      77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;

QY  12 PPSLSVSSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSDGDMWTV 71
DB  40 PGPNSHNS---NTPGIREAGSEDIIVVALDYEAIHEDLSFKQGDQMVVLEESGEWKA 96
QY  72 LSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIRESQTRR 127
DB  97 RSLATRKEGYPISNYVARYVDSLETEEFKGISRKAERQALLAPGNLGSFMRDSEITK 156
QY  128 GYSLSVRLSRPASWDRIRHYRICHLDNGWLYISPRITFPPSLQALVDHYSELADICCLL 187
DB  157 GYSLSVRYDPRQGVTVKHYSRLDNGGYISPRITFPPSLQALVDHYKKGNDGLCQKL 216
QY  188 KEPCV 192
DB  217 SVPCM 221

RESULT 14
US-09-977-260-17
; Sequence 17, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-260-17

Query Match      26.6%; Score 360.5; DB 9; Length 505;
Best Local Similarity 41.6%; Pred. No. 9.6e-27;
Matches      77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;

QY  12 PPSLSVSSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSDGDMWTV 71
DB  40 PGPNSHNS---NTPGIREAGSEDIIVVALDYEAIHEDLSFKQGDQMVVLEESGEWKA 96
QY  72 LSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIRESQTRR 127
DB  97 RSLATRKEGYPISNYVARYVDSLETEEFKGISRKAERQALLAPGNLGSFMRDSEITK 156
QY  128 GYSLSVRLSRPASWDRIRHYRICHLDNGWLYISPRITFPPSLQALVDHYSELADICCLL 187
DB  157 GYSLSVRYDPRQGVTVKHYSRLDNGGYISPRITFPPSLQALVDHYKKGNDGLCQKL 216
QY  188 KEPCV 192
DB  217 SVPCM 221
```

Db 97 RSLATKEGYIPSNYVARVDSLETEEFKISRKAERQLLAPGNLGSFMRDSEYTK 156  
Qy 128 GSYSLSVLSRPASWDRIRHYRHHCLDNGWLYISPLTFPSLQALVDHYSELADICLL 187  
Db 157 GSYSLSVRDYDPRQGDVYKHVKIRTLDNGGFYISPRSTFTLQELVDHYKKGNDGLCQKL 216  
Qy 188 KEPCV 192  
Db 217 SVPCM 221

RESULT 15

US-09-977-261-17  
; Sequence 17, Application US/09977261  
; Publication No. US20030054527A1  
; GENERAL INFORMATION:  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: SURES, IRVINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1259  
; CURRENT APPLICATION NUMBER: US/09/977,261  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-261-17

Query Match 26.6%; Score 360.5; DB 10; Length 505;  
Best Local Similarity 41.6%; Pred No. 9.6e-27;  
Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
Qy 12 PSPSLSSVQGGPVVWEAERSKATAVALGSFAGGPAELSLRLGEPLTIVSEDGWVTV 71  
Db 40 PGPNSHNS--NTPGIREAGSDDIIVVALYDYEAIHHEDLSFQKGDQMVVLEESGEWKA 96  
Qy 72 LSEVSGREYNTIPSVHGVK---SHGWLVECLSRKAEELLLPGNPGGAFLIRESOTRR 127  
Db 97 RSLATKEGYIPSNYVARVDSLETEEFKISRKAERQLLAPGNLGSFMRDSEYTK 156  
Qy 128 GSYSLSVLSRPASWDRIRHYRHHCLDNGWLYISPLTFPSLQALVDHYSELADICLL 187  
Db 157 GSYSLSVRDYDPRQGDVYKHVKIRTLDNGGFYISPRSTFTLQELVDHYKKGNDGLCQKL 216  
Qy 188 KEPCV 192  
Db 217 SVPCM 221

Search completed: February 20, 2004, 09:50:16  
Job time : 594 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:18:27 ; Search time 83 Seconds  
(without alignments)  
992.172 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MSLPERRRLSPSLSSV.....RESLSFYISLNDEAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1028	76.0	259	11	Q7TMX7 mus musculu
2	583	43.1	266	13	Q7T0U4
3	467.5	34.6	282	13	Q7T3X8
4	375	27.7	491	11	Q8C3I0
5	366.5	27.1	502	13	Q9DBK6
6	361.5	26.7	488	13	Q13064
7	352.5	26.1	508	4	Q961N1
8	344.5	25.5	499	11	Q8K2M8
9	340	25.1	509	6	Q95M32
10	338.5	25.0	509	11	Q91X65
11	328	24.2	509	6	Q95KR7
12	320.5	23.7	519	13	Q7ZYH5
13	315.5	23.3	537	11	Q62844
14	313.5	23.2	496	13	Q93411
15	309.5	22.9	541	11	Q99PW1
16	309	22.8	812	15	Q85466

17	307.5	22.7	537	13	Q7ZYK3
18	307	22.7	517	11	Q63206
19	306.5	22.7	541	11	Q8CBP1
20	306	22.6	534	4	Q16248
21	306	22.6	534	6	Q95W31
22	306	22.6	534	11	Q8K2A3
23	305.5	22.6	525	13	Q8AWF1
24	305	22.5	527	5	Q9Y121
25	304.5	22.5	517	5	Q77050
26	301	22.2	517	11	Q8BGM0
27	298.5	22.1	541	11	Q8C762
28	288.5	21.3	587	15	Q64817
29	285	21.1	511	5	Q8WQM5
30	284.5	21.0	503	5	Q8WSU5
31	283	20.9	502	13	Q8QGJ9
32	280.5	20.7	523	15	Q85477
33	280.5	20.7	526	15	Q64994
34	277.5	20.5	482	4	Q8NSD7
35	277.5	20.5	526	15	Q93080
36	276.5	20.4	535	11	Q80XU2
37	276.5	20.4	570	4	Q86VB9
38	276	20.4	535	15	Q92957
39	275.5	20.4	526	15	Q64993
40	275	20.3	546	15	Q86363
41	273.5	20.2	526	15	Q92806
42	273.5	20.2	542	11	Q93J10
43	273	20.2	504	5	Q8WSU2
44	273	20.2	545	15	Q86362
45	272	20.1	537	13	Q7ZX73

#### ALIGNMENTS

#### RESULT 1

Q7TMX7 PRELIMINARY; PRT; 259 AA.

AC Q7TMX7;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6NCr; TISSUE=Hematopoietic Stem Cell;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatane P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6NCr; TISSUE=Hematopoietic Stem Cell;

RA Strausberg R.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBSJ databases.  
 DR EMBL; BC053655; AAH5655.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 259 AA; 28495 MW; F078D1FBB720B7CE CRC64;

Query Match 76.0%; Score 1028; DB 11; Length 259;  
 Best Local Similarity 79.4%; Pred. No. 1.8e-84;  
 Matches 208; Conservative 16; Mismatches 34; Indels 4; Gaps 3;

QY 1 MGSIPRRKSLSPSSSSVQCGPVTWEAERSKATATAVALGSPFAGGPAELSLRLGEPLT 60  
 DB 1 MGSLSRGHT-SPPSSSGPPQEPVSPMPERRKVTAVAGLFPAGEQARSLRLGEPLT 59  
 QY 61 IVSEDDGDMTIVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKABELLLLPNPGAFLLI 120  
 DB 60 IIISEDGDMTIVQSEVSGREYHPSVYVAVKVAHGWLYEGLSREKABELLLLPNPGAFLLI 119  
 QY 121 RESQTRGYSYLSVRLSRPASWDRIHRYHICLDNGWLVIISPLTFPSLQALVDHYSEL 180  
 DB 120 RESQTRGYSYLSVRLSRPASWDRIHRYHICLDNGWLVIISPLTFPSLQALVDHYSEL 179  
 QY 181 DITCCLLKEPCVLQRAQPLPGKIDIPVTVQRTPLNWKELDSLLFSEA-ATGEESLLSE 239  
 DB 180 DGICCPRLPCVLQKLGPLEGKDTPTPTVPTSSLNKKLDRSLFLFAPASGEASLLSE 239  
 QY 240 GLRESLSFYISLNDVAVSLDDA 261  
 DB 240 GLRESLSFYISLNDVAVSLDDA 259

RESULT 2  
 Q7T0U4 ID Q7T0U4 PRELIMINARY; PRT; 266 AA.  
 AC Q7T0U4;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative."  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Ioshizuka S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Heiton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RA Klein S., Strausberg R.;  
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBSJ databases.  
 DR EMBL; BC056035; AAH56035.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 266 AA; 30303 MW; EF20A453A4CD673B CRC64;

Query Match 43.1%; Score 583; DB 13; Length 266;  
 Best Local Similarity 47.5%; Pred. No. 2.5e-44;  
 Matches 125; Conservative 48; Mismatches 80; Indels 10; Gaps 5;

QY 1 MGSIPRRKSLSPSSSSVQCGPVTWEAERSKATATAVALGSPFAGGPAELSLRLGEPL 59  
 DB 1 MGNIPCKGYSATEASLTNLEINQ-----AAADKCIFVALYNFPLSGQTDLSHFGEQL 54  
 QY 60 TIIVSEDDGDMTIVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKABELLLLPNPGAFLL 119  
 DB 55 NILSEDDGDMLVKVSASTGQCYLPTNVAVKVRWLYKGINREKABELLMVNSRSGSFL 114  
 QY 120 IRESQTRGYSYLSVRLSRPASWDRIHRYHICLDNGWLVIISPLTFPSLQALVDHYSEL 179  
 DB 115 IRESETRSGSYLTSIRKTNHTSRDSIKHYRIHQLDNGWFIAPRLTFATLQDMVYDYSEV 174  
 QY 180 ADDICLLKEPCVLQRAQPLPGKIDIPVTVQRTPLNWKELDSLLFSEA-ATGEESLL 237  
 DB 175 ADGICCTLTPEVCVQVRV-PNPVTQRPSEPIVVRKPTLNLTWTKLDSMDLNFKNKDDKLNEDCPL 233  
 QY 238 SEGRLSELSFYISLNDVAVSLDD 260  
 DB 234 SLGLREAVSSYMLMTQDSSEMT 256

RESULT 3  
 Q7T3X8 ID Q7T3X8 PRELIMINARY; PRT; 282 AA.  
 AC Q7T3X8;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Src-like adaptor protein.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lebigot I., Gardellin P., Lefebvre L., Beug H., Ghysdael J.,  
 RA Tran Quang C.;  
 RT "Cloning of the avian Src-like adaptor protein cDNA."  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBSJ databases.  
 DR EMBL; AY278230; AAP34689.1; -;  
 SQ SEQUENCE 282 AA; 31988 MW; 85EA8B1D8986D416 CRC64;

Query Match 34.6%; Score 467.5; DB 13; Length 282;  
 Best Local Similarity 41.7%; Pred. No. 7.2e-34;  
 Matches 105; Conservative 38; Mismatches 76; Indels 33; Gaps 5;

QY 27 TWEAERSKA-----TAVA-----LGSFPAGGPAELSLRLGEPLTIVSDDGDMW 69  
 DB 4 TWKTFRASNEGGEETINVASSTQGESDFLAVLYDPSADISQPIFHVGEKLRVLSDEGGWW 63  
 QY 70 TVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKABELLLLPNPGAFLLRESQTRGS 129  
 DB 64 RVHSLTGTGRENIPGKYVAKVYHGWLFEGLSREKABELLQPNTKVGSFMIRESTRGL 123  
 QY 130 YSLSVRLSRPASWDRIHRYHICLDNGWLVIISPLTFPSLQALVDHYSELADDCILKE 189



```

Db 45 STGPSVFSDG-----ESIALYVDYEGINEGDLGFKKGDKLKILOESGWWRAQSI 96
Qy 75 VSGREYNIPSVHVKVS---HGWLVEGLSRKAEELLIPGPGAGFLIRSEOTRRGSYS 131
Db 97 STGEGIPSNYVAIDSLTEEFKVSRSKDAERQLASGNKMGSMIRDSSTTKGSYS 156
Qy 132 LSVLSPASWDRIHRVIRHCLDNGWLYISPRITFPSLQALVDHYSELADDDICLLKBP 191
Db 157 LSVRDSQSGDVTVKYKIRTLONGFYISPRITFTLQELVSHYKLGDLGCLQALTSFC 216
Qy 192 VLQAGPLPGK 202
Db 217 L-----SPKPK 223

RESULT 6
Q13064
ID O13064 PRELIMINARY; PRT; 488 AA.
AC O13064
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE LYN protein tyrosine kinase.
GN LYN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Fukami Y., Funabiki K., Sato K.;
RT "Nucleotide sequence of Xenopus Lyn protein tyrosine kinase.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC 1. SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AB003358; BAA20078.1; -.
DR HSSP; P08631; IAD5.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:intracellular signaling cascade; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000093; SH2_1.
DR ProDom; PD000066; SH3_1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRKG; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW ATP-binding; Kinase; SH3 domain; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 488 AA; 55794 MW; E7E7068B86EA92B2 CRC64;

Query Match 26.7%; Score 361.5; DB 13; Length 488;
Best Local Similarity 38.3%; Pred. No. 5.6e-24;
Matches 82; Conservative 79; Mismatches 79; Indels 17; Gaps 4;

Qy 1 MSLSPRR-----KSLPSPSLSSVQGGPVTMEARSKATAVALGSPAGPAELS 52

```

```

Db 1 MGCISKTDNPRDANGKNPASRTPSLLPGQ-KMIQDIEEGNIVIALYPOGIHEDDLS 59
Qy 53 LRLGEPLTIVSGDWWTVLSEVSGREYNIPSVHVKV-----SHGWLVEGLSRKAEELL 108
Db 60 FKKGELKVLBEHGEWNAKSLSTKKEGFIPTSNVAVRNTLETEWFFKOLTRKDAERQL 119
Qy 109 LLPNGPGAFITRESOTRRGSYSLSVLSRSPASWDRIHRVIRHCLDNGWLYISPRITFPS 168
Db 120 LAPGNNPQAFIRSETSKGSYSLSIRDCDPTQGVKIKYKIRTLONGGYISPRITFTS 179
Qy 169 LQALVDHYSELADDDICLLKBPVQLQAGPLPGK 202
Db 180 INEMIQHYQKQADGLCRKLDKFCF-----SPKPK 209

RESULT 7
Q96IN1
ID Q96IN1 PRELIMINARY; PRT; 505 AA.
AC Q96IN1
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein (B lymphoid tyrosine kinase).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RT TISSUE=Blood, and Lymph;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007371; AAH07371.1; -.
DR EMBL; BC032413; AAH32413.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:intracellular signaling cascade; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000093; SH2_1.
DR ProDom; PD000066; SH3_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW Hypothetical protein; ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 505 AA; 57706 MW; B5F739BEF8389176 CRC64;

Query Match 26.1%; Score 352.5; DB 4; Length 505;
Best Local Similarity 43.6%; Pred. No. 3.8e-23;
Matches 75; Conservative 24; Mismatches 68; Indels 5; Gaps 2;

Qy 25 PVTMEARSKATAVALGSPAGPAELSRLGLCEPLTIVSGDWWTVLSEVSGREYNIPS 84
Db 51 PPEHLDEDEKGFVALDYDTAMNDRDLQMLKGEKLQVLKGTGDWLARSLLVTGREGVPS 110
Qy 85 VHVYKVS----HGWLVEGLSRKAEELLIPGPGAGFLIRSEOTRRGSYSLSVLSRPA 140
Db 111 NFVARVESLEMERWFRSQKAEARQLLAPINKAGSLIRSETNKGAFLSVK-DVTT 169
Qy 141 SWDIRIRHICLDNGWLYISPRITFPSLQALVDHYSELADDDICLLKBPV 192

```

	Query Match	25.18;	Score 340;	DB 6;	Length 509;
	Best Local Similarity	40.64;	Prod No. 5.1e-22;		
	Matches 73;	Conservative	26;	Mismatches 71;	Indels 10;
				Gaps	2;
Qy	26	VYWEAERSKAT	-----AVALGPPGAPPAELSLRLGSPETIVSDGDVWTVLSEVSGRE	79	
Db	49	VYVEGSPNPASP	LDNLVIALHLYSEPSHDGDLGFEKGEQLRLQSGEMWKAQSLTTSOE	108	
Qy	80	YNIPSVHVGKVS	---HGMVYEGLSREKAEELLILPGNPGGAFILIRSGOTRGVSYLSVR	135	

Db 109 GFIPNFVAKANSLEPEPFFKNLSKDAERQQLAPGNTGSHFLRESSTAGSFSLSVR 168  
 Qy 136 LSRPASWDRRHRYHICLDNGWLYISRLTFPSLQALVDHYSLELADICLLKEPCVLQR 195  
 Db 169 DFDQNGQEVVKIRNLNDNGVYISPRITFPGLHVLRYHTNASDGLTRLRSPQOTQK 228

RESULT 10  
 Q91X65 PRELIMINARY; PRT; 509 AA.  
 AC Q91X65;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to lymphocyte-specific protein tyrosine kinase (2 days neonate  
 thymus thymic cells cDNA, RIKEN full-length enriched library,  
 DE clone:E43002B06 product:lymphocyte protein tyrosine kinase, full  
 DE insert sequence).  
 GN LCK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RA Strausberg B.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hitamoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Taya T., Yasunishi A., Yamatsuta M., Hayashizaki Y.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RA The FANTOM Consortium;  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs."  
 RL Nature 420:563-573 (2002).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690 (2001).  
 [5]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44 (1999).  
 [6]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630 (2000).

RN SEQUENCE FROM N.A.  
 RP STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagata S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Togawa Y., Iizawa M., Chara E., Watanabe M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer."  
 RL Genome Res. 10:1757-1771 (2000).  
 DR EMBL; BC011474; AAH11474.1;  
 DR EMBL; AK088001; BAC40086.1;  
 DR PIR; I48845; I48845;  
 DR MGD; MGI:96756; Lck.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH2DOMAIN.  
 DR PRINTS; PR00103; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00226; SH3; 1.  
 DR SMART; SM00219; TyroK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00001; SH2; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 509 AA; 57942 MW; 3513102F49A7FD0B CRC64;  
 Query Match 25.0%; Score 338.5; DB 11; Length 509;  
 Best Local Similarity 39.1%; Pred. No. 7e-22;  
 Matches 77; Conservative 28; Mismatches 85; Indels 7; Gaps 2;  
 QY 3 SLPSRKSLPSLSVSSVQGGQPVMTAEASRSKATAVALGSPAGGPAELSLRLGPELTV 62  
 DB 35 SLFIRNGSEVRDPL---VTYEGSLPPASPLQDNLVIALHSYEPHSDGDLGPEKGEQLIL 91  
 QY 63 SDGQGWTVLSEVSGREYNIPSVHYGKVS----HGWLYEGLSRKAEELLLLPGNPGGAF 118  
 DB 92 EQSGEWKKAQSLTTGEGFIPNFVAKANSLEPEPFFKNLSRKAERQQLAPGNTGSGF 151  
 QY 119 LRESQTRGYSLSVRLSRPASWDRIRHYRHICLDNGWLYISPELTSPSLQALVDHYSE 178  
 DB 152 LIRESSESTAGSFSLSVRDFDQNGGEVWVXHYKIRNLNDNGGFIYSPRITFPGLHDLVRHTN 211  
 QY 179 LADDICLLKEPCVLQR 195  
 DB 212 ASDGLCTKLSRCPQOTQK 228

RESULT 11  
 Q95KR7 PRELIMINARY; PRT; 509 AA.  
 ID Q95KR7  
 AC Q95KR7;

DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Lck tyrosine kinase (Fragment).  
 DE LCK.  
 OS Saimiri sciureus (Common squirrel monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.  
 ON NCBI\_TaxID=9521;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21424508; PubMed=11533187;  
 RA Greve T, Tamguney G, Fleischer B, Fickenscher H, Brooker B.M.;  
 RT "Downregulation of p56lck tyrosine kinase activity in T cells of  
 squirrel monkeys (Saimiri sciureus) correlates with the non-  
 transforming and apathogenic properties of herpesvirus saimiri in its  
 natural host.";  
 RT J. Virol. 75:9252-9261(2001).  
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL: AJ277921; CAC38871.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR000980; SH2.  
 DR InterPro: IPR001452; SH3.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR InterPro: IPR008266; Tyr\_pkinase\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR PRINTS: PRO0401; SH2DOMAIN.  
 DR PRINTS: PRO0452; SH3DOMAIN.  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR ProDom: PD000093; SH2; 1.  
 DR ProDom: PD000066; SH3; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR SMART: SM00326; SH3; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50001; SH2; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 KW ATP-binding; Kinase; SH3 domain; Transferase; Tyrosine-protein kinase.  
 FT NON\_TER 509 509  
 SQ SEQUENCE 509 AA; 58253 MW; 5088C6407D109519 CRC64;  
 Query Match 24.2%; Score 328; DB 6; Length 509;  
 Best Local Similarity 41.4%; Pred. No. 6.2e-21;  
 Matches 67; Conservative 27; Mismatches 64; Indels 4; Gaps 1;  
 QY 38 VALGSPAGGPAELSLRLGEPTITVSEDGDWTVLSEVSGREYNIPSVHVGKVS-----HG 93  
 Db 67 IALHSTPEPHDGLGFEKEGHELRILLEGQENWKAQSLITGQGGVFFNFVAKANSLPEEP 126  
 QY 94 WLYEGLSREKABELLPGNPGGAFLIRESQTRGYSLSVRLSPASWDRIRHYRIHCL 153  
 Db 127 WFFKNLSRKAERQLLAPGNTGHSFLIRESESTAGSFSLSVRDPQNGQEVVKKIRNL 186  
 QY 154 DNGWLYISRLTFPSLQALVDHYSLEADDICCLKEPCVLOR 195  
 Db 187 DNGGFYISPRITFSGLHVLRYHTNASDGLCTRLSRPQCTQK 228  
 RESULT 12  
 Q7ZYH5  
 ID Q7ZYH5 PRELIMINARY; PRT; 519 AA.  
 AC Q7ZYH5  
 DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Similar to Fyn proto-oncogene.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 ON NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Embryo;  
 RA Klein S., Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC043783; AH43783.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_Thr\_pkinase.  
 DR InterPro: IPR001452; SH2.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR InterPro: IPR008266; Tyr\_pkinase\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR PRINTS: PRO0401; SH2DOMAIN.  
 DR PRINTS: PRO0452; SH3DOMAIN.  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR ProDom: PD000093; SH2; 1.  
 DR ProDom: PD000066; SH3; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR SMART: SM00326; SH3; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50001; SH2; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 SQ SEQUENCE 519 AA; 58486 MW; A790E29424AE7037 CRC64;  
 Query Match 23.7%; Score 320.5; DB 13; Length 519;  
 Best Local Similarity 36.2%; Pred. No. 3e-20;  
 Matches 75; Conservative 37; Mismatches 76; Indels 19; Gaps 5;  
 QY 4 LPSRRKSLPSFLSSVQG--QGP-----VTMEARSKATAVALGSPAGGPAELSL 53  
 Db 33 MPFRGTGTHPTFINTGIGVMSGFSRGMGVGTAGA---TLFIALYDYDARTEDDLTF 88  
 QY 54 RLGEPTITV-SEGDWTVLSEVSGREYNIPSVHVGKVS-----SHGWLYEGLSREKABELL 108  
 Db 89 QKGEKPHIINNSEGDWEARSLSSGNTGTPSNYVAPMNSIQSKDFFGKIGKDAERLL 148  
 QY 109 LLPNPGGAFLIRESQTRGYSLSVRLSPASWDRIRHYRIHCLDNGWLYISRLTFPS 168  
 Db 149 LSEGNPRGTGTHPTFINTGIGVMSGFSRGMGVGTAGA---TLFIALYDYDARTEDDLTF 208  
 QY 169 LQALVDHYSLEADDICCLKEPCVLOR 195  
 Db 209 VEELVEHYGFENDGLCHLLTSPCINQK 235  
 RESULT 13  
 Q62844  
 ID Q62844 PRELIMINARY; PRT; 537 AA.  
 AC Q62844  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:34:08 ; Search time 3092 Seconds  
(without alignments)  
3658.645 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLSRKLSPSPSLSSV.....RESLSFVISINDEAVSLDDA 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cn2\_1/uspto.spool/p/US09939853/runat\_19022004\_145339\_24464/app\_query.fasta\_1.455  
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09939853 @CN 1 1 5265 @runat\_19022004\_145339\_24464 -NCPU=6 -ICPU=3  
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

- 1: gb.ba:
- 2: gb.htg:
- 3: gb.in:
- 4: gb.om:
- 5: gb.ov:
- 6: gb.pat:
- 7: gb.ph:
- 8: gb.pl:
- 9: gb.pr:
- 10: gb.ro:
- 11: gb.sts:
- 12: gb.sv:
- 13: gb.un:
- 14: gb.vi:
- 15: em.ba:
- 16: em.fun:
- 17: em.hum:
- 18: em.in:
- 19: em.mu:
- 20: em.om:
- 21: em.or:
- 22: em.ov:
- 23: em.pat:
- 24: em.ph:
- 25: em.pl:
- 26: em.rc:
- 27: em.sts:
- 28: em.un:

- 29: em.vi:
- 30: em.htg.hum:
- 31: em.htg.inv:
- 32: em.htg.other:
- 33: em.htg.mus:
- 34: em.htg.pln:
- 35: em.htg.rtd:
- 36: em.htg.mam:
- 37: em.htg.vrt:
- 38: em.sy:
- 39: em.htgo.hum:
- 40: em.htgo.mus:
- 41: em.htgo.other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1353	100.0	1183	6	AX443133 Sequence
2	1353	100.0	1183	6	AX443135 Sequence
3	1347	99.6	786	6	AX511153 Sequence
4	1347	99.6	786	6	AX572845 Sequence
5	1347	99.6	786	9	AF290985 Homo sapi
6	1347	99.6	786	9	AF326353 Homo sapi
7	1347	99.6	2415	9	AK025645 Homo sapi
8	1347	99.6	2538	9	BC042041 Homo sapi
9	1347	99.6	2567	6	AX452880 Sequence
10	1341	99.1	2788	6	AX780857 Sequence
11	1196.5	88.4	737	6	AX511155 Sequence
12	1196.5	88.4	737	6	AF290986 Homo sapi
13	1028	76.0	777	6	AX511151 Sequence
14	1028	76.0	1348	6	AX511150 Sequence
15	1028	76.0	1348	10	AF287467 Mus muscu
16	1028	76.0	1384	10	BC052655 Mus muscu
17	1019	75.3	1321	10	AF434990 Mus muscu
18	583	43.1	1571	5	BC056035 Xenopus l
19	501	37.0	1631	10	AY217759 Rattus no
20	482	36.4	921	10	MMU29056 Mus muscu
21	482	36.4	2648	10	BC032922 Mus muscu
22	482	36.4	2773	10	MMU131777 Mus muscu
23	491.5	36.3	841	10	AY079449 Mus muscu
24	491.5	36.3	841	10	AY079450 Mus muscu
25	487	36.0	1076	9	HSU30473 Homo sapien
26	487	36.0	1870	9	BC007042 Homo sapi
27	487	36.0	2021	9	HSU44403 Human Src-1
28	487	36.0	2109	6	AX428893 Sequence
29	487	36.0	2665	6	AX333017 Sequence
30	487	36.0	2665	9	D89077 Homo sapien
31	487	36.0	3090	9	HSM808530 Homo sapi
32	482	35.6	145068	9	HSU44403 Human DNA
33	482	35.6	145833	2	AC026539 Homo sapi
34	467.5	34.6	849	5	AY278230 Gallus ga
35	380.5	28.1	2703	10	MUSLYNB Mouse lyn B
36	375.5	27.8	1538	10	RATLYNBYR L14823 Rattus norv
37	375.5	27.8	2302	10	RATLYNBYR AF000301 Rattus no
38	375.5	27.8	2824	10	RNAF000302 Rattus no
39	370.5	27.4	1539	12	AY335782 Synthetic
40	370.5	27.4	2298	6	AR380613 Sequence
41	370.5	27.4	2298	6	AX779985 Sequence
42	370.5	27.4	2298	9	HUMLYN M16038 Human lyn m
43	370.5	27.4	4175	6	AX779984 Sequence
44	370.5	27.4	1601	10	RATLYNBYR L14782 Rattus norv
45	370	27.3	2041	10	RNAF000300 Rattus no

ALIGNMENTS

RESULT 1

```

AX443133
LOCUS AX443133 1183 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 74 from Patent WO0216599.
ACCESSION AX443133
VERSION AX443133.1 GI:21690555
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,
Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,
Topper,J.N. and Yang,R.B.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;
Curation Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES
    source
        1. 1183
            Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2,01e-102 Length: 1183
Score: 1353.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-939-853A-75 (1-261) x AX443133 (1-1183)
QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
DB 398 ATGGAGAGTGTGCCACAGAGAGAGAAATCTCTCCAGGCCCAAGCTTGATTCCTCTGTC 457
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
DB 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGAGCCAGCCGCTGGGCCCTG 517
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCGTGAGACTCGGGAGCCATTGACC 577
QY 61 IleValSerGluAspGlyAspTrrTrpThrValLeuSerGluValSerGlyArgGluTyr 80
DB 578 ATCGTCTCTGAGGATGGAGACTGTGTCGACGCTGTCTGTAAGTCTCAGGCAGAGATAT 637
QY 81 AsnIleProSerValHisValGlyValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
DB 638 AACATCCCCAGCTCCAGTGGGCAAAAGTCTCCATGGTGGCTGTATGAGGCCCTGAGC 697
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120
DB 698 AGGAGAAAGCAGAGGAACCTCTGTGTGTTACTCGGAACCTCGAGGGGCTTCTTCATC 757
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 758 CGGGAGACCCAGACAGAGAGGCTCTTACTCTCTGTCAGTCCCTTCAGCCGCCCTGCA 817
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
DB 818 TCCTGGGACCGATCAGACACTACAGGATCCACTGCCCTTGACATGCTGGCTGTATCATC 877
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
DB 878 TCACGGCCCTCACCTTCCCTCACTCCAGGCCCTGGTGGACCATTAATCTGAGCTGGCG 937
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
DB 938 GATGACATCTGCTACTCAAGGAGGCCCTGTGCTCTGACAGAGGCTGGCCCGCTCCCT 997

```

```

QY 201 GlyIysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
DB 998 GGCAAGGATATACCCCTACCTGTGACTGTGACAGGACACCACTCACTGGAAGAGCTG 1057
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuSerGluGly 240
DB 1058 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTCTCACTGAGGGT 1117
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260
DB 1118 CTCGGGAGTCCCTCAGCTCTTACATCAGCCTGATGACGAGGCTGTCTCTTTGGATGAT 1177
QY 261 Ala 261
DB 1178 GCC 1180
RESULT 2
AX443133/c
LOCUS AX443133 1183 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 76 from Patent WO0216599.
ACCESSION AX443133
VERSION AX443133.1 GI:21690556
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,
Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,
Topper,J.N. and Yang,R.B.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;
Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES
    source
        1. 1183
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2,01e-102 Length: 1183
Score: 1353.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-939-853A-75 (1-261) x AX443133 (1-1183)
QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
DB 786 ATGGAGAGTGTGCCACAGCAGAGAGAAATCTCTCCAGGCCCAAGCTTGATTCCTCTGTC 727
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
DB 726 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGCCAGCCGCTGGCCCTG 667
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB 666 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCGTGAGACTCGGGAGGCCATTGACC 607
QY 61 IleValSerGluAspGlyAspTrrTrpThrValLeuSerGluValSerGlyArgGluTyr 80
DB 606 ATCGTCTCTGAGGATGGAGACTGTGTCGACGCTGTCTGTAAGTCTCAGGCAGAGATAT 547
QY 81 AsnIleProSerValHisValGlyValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
DB 546 AACATCCCCAGGCTCCAGTGGGCAAAAGTCTCCATGGTGGCTGTATGAGGCCCTGAGC 487
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120

```

Db 486 AGGAGAAAGCAGAGGAACCTCTGTGTACCTGGAAACCTGGAGGGCTTCTCTCATC 427  
Qy 121 ArgGluSerGlnThrArgArgGlySerTySerLeuSerValArgLeuSerArgProAla 140  
Db 426 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTACGTCGCCCTCAGCCGCCCTGCA 367  
Qy 141 SerTrpAspArgIleArgHisTyArgIleHisCysLeuAspAsnGlyTrpLeuTyIle 160  
Db 366 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCTTGACATGCTGGCTGTATCATC 307  
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTySerGluLeuAla 180  
Db 306 TCACCGGCCCTCACTTCCCTCTCACTCAGGCCCTGGTGACCACTTACTCTGAGCTGGCG 247  
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
Db 246 GATGACATCTGCTGCTACTCAAGAGGCCCTGTGTCTCGAGAGGCTGGCCGCCCTCCT 187  
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
Db 186 GCGAAGGATATACCCCTACCTGTGACTGTGCAGAGGACACCACTCAACTGGAAAGAGCTG 127  
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
Db 126 CACAGCTCCCTCTGTTTCTGAGCTGCCACAGGGAGGAGTCTCTCTCAGTGAGGT 67  
Qy 241 LeuArgGluSerLeuSerPheTyIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
Db 66 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGGAATGACGAGGCTGTCTCTTTGGATGAT 7  
Qy 261 Ala 261  
Db 6 GCC 4

RESULT 3  
AX511153  
LOCUS AX511153 786 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 4 from Patent WO0242452.  
ACCESSION AX511153  
VERSION AX511153.1 GI:23392046  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS McGlade, J.C. and Loreto, M.P.  
TITLE Adapter gene  
JOURNAL Patent: WO 0242452-A 4 30-MAY-2002;  
The Hospital for Sick Children (CA)  
FEATURES  
source  
1. 786  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 3 9e-102 Length: 786  
Score: 1347.00 Matches: 260  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.56% Indels: 0  
DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x AX511153 (1-786)

Qy 1 MetGlySerLeuProSerArgArgLysSerLeuProProSerProSerLeuSerSerVal 20  
Db 1 ATGGGAAGTCTCCAGAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTCTGTC 60  
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40

Db 61 CAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG 120  
Qy 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGluProLeuThr 60  
Db 121 GGCAGTTTCCCGGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGGCATGACC 180  
Qy 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyI 80  
Db 181 ATCGTCTCTGAGATGGAGACTGGTGGACGGTGTCTGAAGTCTCAGGCAGAGAGTAT 240  
Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyIleGlyLeuSer 100  
Db 241 AACATCCCGCAGGTCCAGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCCCTGAGC 300  
Qy 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120  
Db 301 AGGAGAGAGCAGAGAACTGCTGTGTGTACCTGGGAACCTGGAGGGGCTTCTCTCATC 360  
Qy 121 ArgGluSerGlnThrArgArgGlySerTySerLeuSerValArgLeuSerArgProAla 140  
Db 361 CGGAGAGCAGACAGAGAGAGGCTTACTCTCTGTGAGTCCGCTCAGCCGCCCTGCA 420  
Qy 141 SerTrpAspArgIleArgHisTyArgIleHisCysLeuAspAsnGlyTrpLeuTyIle 160  
Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACAATGGCTGGCTGTACATC 480  
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTySerGluLeuAla 180  
Db 481 TCACCGGCCCTCACCTTCCCTCACTCCAGGCCCTGGTGGACCATTACTCTCAGCTGGCG 540  
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
Db 541 GATGACATCTGCTGCTACTCAAGAGGCCCTGTGTCTGCAGAGGCTGGCCGCCCTCCT 600  
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
Db 601 GCGAAGGATATACCCCTACCTGTGACTGTGCAGAGGACACCACTCAACTGGAAAGAGCTG 660  
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
Db 661 GACAGTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTCTCAGTGAGGT 720  
Qy 241 LeuArgGluSerLeuSerPheTyIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGGAATGACGAGGCTGTCTCTTTGGATGAT 780  
Qy 261 Ala 261  
Db 781 GCC 783

RESULT 4  
AX572845  
LOCUS AX572845 786 bp DNA linear PAT 29-NOV-2002  
DEFINITION Sequence 1 from Patent WO02055707.  
ACCESSION AX572845  
VERSION AX572845.1 GI:26004935  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Holland, S.J., Mendenhall, M.K., Pardo, J., Spencer, C., Fu, A.C.,  
Luo, Y., Payan, D.G., Mancebo, H.S., Wu, J., Zhou, X., Shen, M.,  
Liao, X.C. and Sheng, N.  
TITLE Cloning of an inhibitor of antigen-receptor signaling by a  
retroviral-based functional screen  
JOURNAL Patent: WO 02055707-A 1 18-JUL-2002;  
Rigel Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1. 786  
/organism="Homo sapiens"  
source

```
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..786
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD58542.1"
/db_xref="GI:26004936"
/db_xref="RENTREMBL:CAD58542"
/translation="MGSLSRRKSLSPSLSSVQGGPVVMEARSKATAVALGSPF
AGPAELSLRLGPELTVSESDGDMWTVLSEVSGREYNIPSVHAKVSHGWLYEGLSRE
KAEELLLPQNGAFILRESQTRGSYSLSVRLSRPASMDRIKHYRIHCLDNGWLYI
SPRLTFPSLQALVDHYSELADDICLLKEPCVLQKAGPLPGKDIPLPVTQRTPLNWK
ELDSLLFSEAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"
ORIGIN
Alignment Scores:
Pred. No.: 3.9e-102 Length: 786
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.56% Indels: 0
DB: 6 Gaps: 0
US-09-939-853A-75 (1-261) x AX572845 (1-786)
Qy 1 MetGlySerLeuProSerArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGATTCCTCTGTC 60
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGCCAGGAGACCTGTGACCATGGAAGCAGAGAGAGCAAGCCAGCCGTCGGCCCTG 120
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGluProLeuThr 60
Db 121 GGCAGTTCCTCCGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Qy 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCTCTCTCAGGATGGAGACTGTGTGAGCGGTGTGTCTGAAGTCTCAGCAGAGATAT 240
Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCAGCTCCAGTGGCCGCAAGTCTCCATGGGTGGTGTGTATGAGGCGCTGAGC 300
Qy 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120
Db 301 AGGAGAAAGCAGAGGAACCTGCTGTGTACCTGGGAACCTGGAGGGGCTTCTCATC 360
Qy 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGCAGAGAGGCTTACTCTCTGTGAGTCCGCTCAGCCGCTGCA 420
Qy 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db 421 TCCTGGGACCGATCAGACACTCAGGATCCACGTCCCTGCAATGCTGGCTGTATCATC 480
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAlaPheHisTyrSerGluLeuAla 180
Db 481 TCACCGCGCCCTCACCTCCCTCTACTCCAGGCGCTGTGTGAGCACCATTACTCTGAGCTGGC 540
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 541 GATGACATCTCTGCTACTCAAGAGGCGCTGTGTCTGTGAGAGGGCTGGCCGCTCCCT 600
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 601 GGCAGGATATACCCCTACCTGTGACTGTGACAGAGGACACCACTCAACTGGAAGAGCTG 660
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 661 GACAGTCCCTCTCTGTTTCTGAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGT 720
Qy 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
Db 721 CTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATCAGGAGCTGTCTTTTGGATGAT 780
Qy 261 Ala 261
Db 781 GCC 783
RESULT 5
AF290985 786 bp mRNA linear PRI 21-JAN-2003
LOCUS Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
DEFINITION AF290985
ACCESSION AF290985
VERSION AF290985.1 GI:17351920
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLES Cloning and characterization of human Src-like adaptor protein 2
and a novel splice isoform, SLAP-2-v
JOURNAL Oncogene 22 (2), 266-273 (2003)
MEDLINE 22415750
PUBMED 12527895
REFERENCE 2 (bases 1 to 786)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLES Direct Submission
JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
source
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1..786
/notes="SLAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAL38197.1"
/db_xref="GI:17351921"
/translation="MGSLSRRKSLSPSLSSVQGGPVVMEARSKATAVALGSPF
AGPAELSLRLGPELTVSESDGDMWTVLSEVSGREYNIPSVHAKVSHGWLYEGLSRE
KAEELLLPQNGAFILRESQTRGSYSLSVRLSRPASMDRIKHYRIHCLDNGWLYI
SPRLTFPSLQALVDHYSELADDICLLKEPCVLQKAGPLPGKDIPLPVTQRTPLNWK
ELDSLLFSEAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"
CDS
Qy 3.9e-102 Length: 786
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.56% Indels: 0
DB: 9 Gaps: 0
US-09-939-853A-75 (1-261) x AF290985 (1-786)
Qy 1 MetGlySerLeuProSerArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGATTCCTCTGTC 60
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGCCAGGAGACCTGTGACCATGGAAGCAGAGAGAGCAAGCCAGCCGTCGGCCCTG 120
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGluProLeuThr 60
Db 121 GGCAGTTCCTCCGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
```

QY 61 lleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 181 ATCGTCTCTGAGGATGGAGATGGTGGACGGTGTCTGTAAGTCTCAGGACAGAGATAT 240  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTrpGluGlyLeuSer 100  
 Db 241 AACATCCCCAGCGTCCACGTGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLe 120  
 Db 301 AGGGAGAAAGCAGAGAACTGCTGTGTACTCGGAAACCCCTGGAGGGCTTCTCTCATC 360  
 QY 121 ArgGluSerGlnThrArgGlySerTrpSerLeuSerValArgLeuSerArgProAla 140  
 Db 361 CGGAGAGCCAGACCCAGAGAGGCTTCTTACTCTGTCTGAGTCCGCTCAGCGCCCTGCA 420  
 QY 141 SerTrpAspArgIleArgHisTrpArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCTTGACAATGGCTGGCTGTATC 480  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTrpSerGluLeuAla 180  
 Db 481 TCACCGCGCTCACCTTCCCTCACTCCAGCCCTGGTGGACCATTAATCTGAGCTGGCG 540  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 Db 541 GATGACATCTGTGCTACTCAAGAGAGCCCTGTGTCTGAGAGGGCTGGCCCGCTCCCT 600  
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 Db 601 GGCAGGATATACCTTACCTTGTGACTGTGAGAGGACACCACTCACTGGAAAGAGCTG 660  
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGluGly 240  
 Db 661 GACAGCTCCCTCTCTTCTTCTGAGCTGCCACAGGGAGAGTCTCTTCTCAGTGAGGGT 720  
 QY 241 LeuArgGluSerLeuSerPheTrpIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
 Db 721 CTTCCGGAGTCCCTCAGCTTCTACATCAGCTGAAATGAGAGGCTGTCTCTTGTGATGAT 780  
 QY 261 Ala 261  
 Db 781 GCC 783  
 RESULT 6  
 AF326353 786 bp mRNA linear PRI 08-NOV-2001  
 LOCUS Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.  
 DEFINITION  
 ACCESSION AF326353  
 VERSION AF326353.1 GI:16797891  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 786)  
 AUTHORS Holland,S.J., Liao,X.C., Mendenhall,M.K., Zhou,X., Pardo,J.,  
 Chu,P., Spencer,C., Fu,A.C., Sheng,N., Yu,P., Pali,E., Nagin,A.,  
 Shen,M., Yu,S., Chan,S., Wu,X., Li,C., Woiseschlager,M., Y.,  
 Aversa,G., Kolbinger,F., Bennett,W.K., Molineaux,S., Luo,Y.,  
 Payan,D.G., Mancebo,H.S.Y. and Wu,J.  
 TITLE Functional cloning of Src-like Adaptor Protein-2 (SLAP-2), a Novel  
 Inhibitor of Antigen Receptor Signaling  
 J. Exp. Med. 194 (9), 1263-1276 (2001)  
 JOURNAL 21553259  
 MEDLINE 1169592  
 PUBMED 2 (bases 1 to 786)  
 REFERENCE Holland,S.J., Mendenhall,M.K., Zhou,X., Spencer,C., Pardo,J.,  
 Fu,A.C., Sheng,N., Shen,M., Liao,C., Luo,Y., Payan,D.G.,  
 Mancebo,H.S.Y. and Wu,J.  
 AUTHORS Direct Submission  
 TITLE Submitted (05-DEC-2000) Rigel Pharmaceutical Inc., 240 East Grand  
 JOURNAL Avenue, South San Francisco, CA 94080, USA

FEATURES  
source

Location/Qualifiers  
 1..786  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"

## CDS

1..786  
 /note="SLAP2; SH2 and SH3 domain-containing adaptor"  
 /codon\_start=1  
 /product="Src-like adapter protein-2"  
 /protein\_id="AAL29204.1"  
 /db\_xref="GI:16797892"  
 /translation="MGSLPSRRKSLPSLSSSVQGGPVTWEERSKATAVALGSPF  
 AGPRLSLRGLPTIVSDGDMTWLSEVSGREYNIPSVHAKVSHGWLVEGJSRE  
 KAEILLPLNPGGAFILRESOTRRGSYLSVRLSRPASWDRIHYRIHCLDNGWLYI  
 SRLTFPSLQALVDHYSELADIDICLLKPCVLQRAGPLGKDIPLPVTQRTPLNWK  
 ELDSLLIFSEATGESLSEGLRESLSFYISLNDEAVSLDDA"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,9e-102 Length: 786  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.56% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-939-853A-75 (1-261) x AF326353 (1-786)  
 QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 Db 1 ATGGGAGTCTGCCACAGCAGAGAAATCTTGCCACAGCCCAAGCTTGAGTTCTCTGTC 60  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 61 CAAGGCGCAGGACCTGTGACCATGGAAAGCAGAGAGAGCAAGCCACACCGCGGCCCTG 120  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 121 GGCAGTTTCCCGCAGGTGGCCCGCGCGAGCTGTGCTGAGACTCGGGAGGCCATTGACC 180  
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 181 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTCTGAAAGTCTTCAGGACAGAGAT 240  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTrpGluGlyLeuSer 100  
 Db 241 AACATCCCGAGCTCCACGTGGCCAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGC 300  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLe 120  
 Db 301 AGGGAGAAAGCAGAGAACTGCTGTGTGTACTCGGAAACCCCTGGAGGGCTTCTCTCATC 360  
 QY 121 ArgGluSerGlnThrArgArgGlySerTrpSerLeuSerValArgLeuSerArgProAla 140  
 Db 361 CGGAGAGCCAGACCCAGAGAGGCTTCTTACTCTGTCTGAGTCCGCTCAGCGCCCTGCA 420  
 QY 141 SerTrpAspArgIleArgHisTrpArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCTTGACAATGGCTGGCTGTATC 480  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTrpSerGluLeuAla 180  
 Db 481 TCACCGCGCTCACCTTCCCTCACTCCAGGAGCCCTGTGCTTCGAGAGGCTGGCCCGCTCCCT 600  
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 Db 601 GGCAGGATATACCTTACCTTGTGACTGTGAGAGGACACCACTCACTGGAAAGAGCTG 660

QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
 Db 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTTCTCAAGTGAAGGT 720

QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
 Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGAAATCAGCAGGAGCTGCTCTTTGGATGAT 780

QY 261 Ala 261  
 Db 781 GCC 783

RESULT 7  
 AK025645 2415 bp mRNA linear PRI 13-SEP-2003  
 LOCUS Homo sapiens cDNA: FL21992 fis, clone HEP06554.  
 DEFINITION  
 ACCESSION AK025645  
 VERSION AK025645.1 GI:10438227  
 KEYWORDS oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 AUTHORS Kawabata,A., Hiki,j,i,T., Kobatake,N., Inagaki,H., Ikema,Y.,  
 Okamoto,S., Ohtani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,  
 Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.  
 NEDO human cDNA sequencing project  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 2415)  
 AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,  
 Shibahara,T., Tanaka,T. and Nakamura,Y.  
 Direct Submission  
 JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,  
 University of Tokyo, Laboratory of Genome Structure Analysis, Human  
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,  
 Japan (E-mail:flcdna@ms.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
 Fax:81-3-5449-5416)  
 COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan; cDNA full insert  
 sequencing; Research Association for Biotechnology; cDNA library  
 construction, 5'- & 3'-end one pass sequencing; Department of  
 Virology and Human Genome Center, Institute of Medical Science,  
 University of Tokyo (partly supported by Science and Technology  
 Agency).

FEATURES source  
 1..2415  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="HEP06554"  
 /cell\_lines="HepG2"  
 /cell\_type="hepatoma"  
 /clone\_lib="HEP"  
 /note="Cloning vector pME18SFL3"  
 61..846  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAB15201.1"  
 /db\_xref="GI:10438228"  
 /translation="MGSLPGRKSLSPSP:SSSVQGGPVTMEABRSKATAVALGSPFP  
 AGPAELSLRIGLEPLTVSDGDWMTVLSEVSGREYNIPSHVAKYSHGWLYEGLSRE  
 KAELLLLPNPGGAFLIRISQTRGSYSLSVRLSPASWDRIHRHRIHCLDNGWLYI  
 SPRTFFPSLQALVDHYSELADIDICLLKEPCVLQRAGLPKDIPVTVTQPTPLNWK  
 ELDSLFLFSEATGEESLLSGLSELSFYSISLNDEAVSLDDA"

CDS  
 61..846  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAB15201.1"  
 /db\_xref="GI:10438228"  
 /translation="MGSLPGRKSLSPSP:SSSVQGGPVTMEABRSKATAVALGSPFP  
 AGPAELSLRIGLEPLTVSDGDWMTVLSEVSGREYNIPSHVAKYSHGWLYEGLSRE  
 KAELLLLPNPGGAFLIRISQTRGSYSLSVRLSPASWDRIHRHRIHCLDNGWLYI  
 SPRTFFPSLQALVDHYSELADIDICLLKEPCVLQRAGLPKDIPVTVTQPTPLNWK  
 ELDSLFLFSEATGEESLLSGLSELSFYSISLNDEAVSLDDA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.44e-101 Length: 2415  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatch: 1  
 Query Match: 99.56% Indels: 0

DB: 9 Gaps: 0  
 US-09-939-853a-75 (1-261) x AK025645 (1-2415)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 61 ATGGGAAGTCTGCCAGCAGAGAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 120

QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluSerIysAlaThrAlaValAlaLeu 40  
 Db 121 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCACAGCCGCTGGCCCTG 180

QY 41 GlySerPheProAlaGlyProAlaGluLeuSerLeuArgGlyGluProLeuThr 60  
 Db 181 GCGAGTTTCCCGCAGCGTGGCCCGCGAGCTGTGCTGAGACTCGGAGGCCATTGACC 240

QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyValGluTyr 80  
 Db 241 ATCGTCTCTGAGGATGGAGACTGTGTGGACGTGTCTGAAGTCTCAGGAGAGAGTAT 300

QY 81 AsnIleProSerValHisValGlyIysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
 Db 301 AACATCCCGCAGCTCCAGTGGCCAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGC 360

QY 101 ArgGluIysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120  
 Db 361 AGGGAGAAGCAGAGGAACTGCTGTGTGTACTCTGGAAACCTCGAGGGGCTTCTCTCATC 420

QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 421 CGGAGAGCCAGACCCAGGAGAGGCTCTTACTCTCTGTCTGCTCCGCTCAGCGGCCCTGCA 480

QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160  
 Db 481 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCTTGACATGGCTGGCTGTACATC 540

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 541 TCACCGCGCTCACCCTTCCCTCCTCCTCCAGGCCCTGGTGGACCATTAATCTTGAGCTGGCG 600

QY 181 AspAspIleCysCysLeuLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 Db 601 GATGACATCTGTGCTACTCAAGGAGCCCTGTGTCTTCAGAGGGCTGGCCCGCTCCCT 660

QY 201 GlyIysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrIysGluLeu 220  
 Db 661 GGCAGAGGATATACCCCTACCTGTGACTGTGAGAGGACCCACTCACTTGAAGAGCTG 720

QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGluGly 240  
 Db 721 GACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGT 780

QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
 Db 781 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGAAATGAGCAGGAGCTGCTCTTTGGATGAT 840

QY 261 Ala 261  
 Db 841 GCC 843

RESULT 8  
 BC042041 2538 bp mRNA linear PRI 07-OCT-2003  
 LOCUS Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (cDNA  
 clone MGC:49845 IMAGE:4429896), complete cds.  
 DEFINITION  
 ACCESSION BC042041  
 VERSION BC042041.1 GI:27469842  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2538)

## AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toehiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bobak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butcherfield, Y.S., Krzywinski, M.I., Skalska, J., Smalios, D.E., Schnerch, A., Schein, J.B., Jones, S.J. and Marra, M.A.

**TITLE**  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)

**MEDLINE**  
22388257

**PUBMED**  
12477932

**REFERENCE**  
2 (bases 1 to 2538)

**AUTHORS**  
Strausberg, R.

**TITLE**  
Direct Submission

**JOURNAL**  
Submitted (23-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

**REMARK**  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

**COMMENT**  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
cDNA Library preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-sngc.stanford.edu>  
Contact: (Dickson, Mark) [mcdepaxil.stanford.edu](mailto:mcdepaxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 88 Row: a Column: 20  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28416422.

## FEATURES

## source

Location/Qualifiers  
1..2538  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:49845 IMAGE:4429896"  
/tissue type="Prostate, adenocarcinoma."  
/clone.lib="NIH MGC\_91"  
/lab host="DH10B"  
/note="vector: PCMV-SPORT6"  
1..2538  
/gene="SLA2"  
/note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845, SLAP2"  
/db\_xref="LocusID:84174"  
/db\_xref="MIM:606577"

## gene

## CDS

363..1148  
/codon\_start=1  
/product="Src-like adaptor 2, isoform a"  
/protein\_id="AAH42041.1"  
/db\_xref="GI:27469843"  
/db\_xref="LocusID:84174"  
/translation="MGSLSPPSRKSLPSPSSVQGGPVTMEARSKATAVALGSGFP  
AGGPALSLRLGLEPTIVSEDDWTVLSEGREYNI PSVHAKVSHGWLVEGLSRE  
KAEELLLPGNPGAFILRESQTRGYSLSVRLSRPASMDRIHRYHCLDNGWLYI

## misc\_feature

SPRLTFPSLQALVDHYSELADDICLLKEPCVLQAGLPCKDIPLPVTVORTPLNWK  
ELDSSLFSEATGEESLLSEGLRESUSFYISLNDEAVSLDDA"  
468..632  
/note="SH3; Region: SH3 domain. SH3 (src homology 3)  
domains are often indicative of a protein involved in  
signal transduction related to cytoskeletal organization.  
First described in the Src cytoplasmic tyrosine kinase.  
The structure is a partly opened beta barrel"  
/db\_xref="CDD:pfam00018"  
642..890  
/note="SH2; Region: SH2 domain"  
/db\_xref="CDD:pfam00017"

## misc\_feature

642..890  
/note="SH2; Region: SH2 domain"  
/db\_xref="CDD:pfam00017"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,538-101 Length: 2538  
Score: 1347.00 Matches: 260  
Percent Similarity: 99.62% Conservatives: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.56% Indels: 0  
DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x BC042041 (1-2538)

QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
DB	363	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTCTC	422
QY	21	GingGlyGingLysProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
DB	423	CAAGCCAGGGACCTGTACCATGGAGCAGAGACAGACAGCCAGCCGTGGCCCTG	482
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGlyProLeuThr	60
DB	483	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCTGCTGAGACTCGGGAGCCATTGACC	542
QY	61	IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr	80
DB	543	ATCGTCTCTGAGATGGAGACTGGTGGACGGTGTCTGAACTCTCAGGCAGAGAGTAT	602
QY	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
DB	603	AACATCCCGCAGGCTCCAGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCCTGAGC	662
QY	101	ArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyValaPheLeuIle	120
DB	663	AGGGAGAAAGCAGAGAACTGCTGTGTGTACTGGGAACCTTGGAGGGCCCTTCTCATC	722
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
DB	723	CGGGAGAGCCAGCCAGGAGAGGCTTACTCTCTGTCTCAGTCCGCTCAGCGCCCTGCA	782
QY	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
DB	783	TCTGGGACCGGATCAGACACTACAGGATCCACTCCCTTGACAAATGGCTGGCTGACATC	842
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
DB	843	TACCGGGCTTCACTTCCCTCACTCCAGGCCCTGGTGGACCATTAATCTGAGCTGGCG	902
QY	181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
DB	903	GATGACATCTGCTGCTACTCAAGGAGGCCCTGTGCTTCGACAGAGGGCTGGCCGCTCCCT	962
QY	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
DB	963	GGCAGGATATACCTTACCTGTGACTGTGCAGAGGACACCACTCACTGGAAGAGCTG	1022
QY	221	AspSerSerLeuLeuPheSerGluAlaThrGlyGluGluSerLeuLeuSerGluGly	240
DB	1023	GACAGTCCCTCTCTGTTTCTGAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGT	1082
QY	241	LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp	260



Db 1083 CTCGGGAGTCCCTCAGCTTCTACATGAGCTGAATGAGAGGCTGCTCTTTGGATGAT 1142  
Qy 261 Ala 261  
Db 1143 GCC 1145  
RESULT 9  
AX452880 2567 bp DNA linear PAT 06-JUL-2002  
LOCUS Sequence 1 from Patent WO0242457.  
DEFINITION AX452880  
ACCESSION AX452880  
VERSION AX452880.1 GI:21712520  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Chang H., Yang W.P., Wu Y., Whitney G.S., Perez-Villar J.J. and  
Kanner S.B.  
TITLE Cloning and expression of human slap-2: a novel sh2/sh3  
domain-containing human slap homologue having immune cell-specific  
expression  
JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;  
Bristol-Myers Squibb Co. (US)  
FEATURES  
source  
Location/Qualifiers  
1..2567  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores: 1.55e-101 Length: 2567  
Pred. No.: 1347.00 Matches: 260  
Score: 99.62% Conservative: 0  
Percent Similarity: 99.62% Mismatches: 1  
Best Local Similarity: 99.62% Indels: 0  
Query Match: 99.56% Gaps: 0  
DB: 6  
US-09-939-853A-75 (1-261) x AX452880 (1-2567)  
Qy 1 MetGlySerLeuProSerArgAlaGlySerLeuProSerProSerLeuSerSerVal 20  
Db 415 ATGGAGTCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGATTCCTCTGTC 474  
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
Db 475 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 534  
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
Db 535 GGCAGTTTCCCGCAGGATGGCCCGCCGAGCTGCGCTGACATCGGGAGCCATTGACC 594  
Qy 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
Db 595 ATCGTCTCTAGGATGGAGACTGTGGACGCTGTGCTGAAGTCTCAGGCAGAGACTAT 654  
Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
Db 655 AACATCCCGCAGCTCCAGTGGCCCAAGATCTCCCATGGTGGCTGTATGAGGCGCTGAGC 714  
Qy 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120  
Db 715 AGGAGAAAGCAGAGGAATCTGTGTATTACCTGGGAACCTCGAGGGGCTTCTCATC 774  
Qy 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
Db 775 CGGGAGAGCCAGCAGAGAGGCTCTTACTCTCTGTGAGTCCGCCCTCAGCCGCCCTGCA 834  
Qy 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160

Db 835 TCCTGGGACCGATCAGACTACAGGATCCATGCTTGCATATGGCTGGCTGATCATC 894  
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
Db 895 TCACGGCGCTTCACCTTCCCTCCTCAGGCCCTGGTGACCATTTACTCTGAGCTGGCG 954  
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
Db 955 GATGACATCTGCTGCTTACTCAGAGGCCCTGTGTCTCGAGAGGCTGGCCGCTCCCT 1014  
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
Db 1015 GGCAGGATATACCCCTACCTGTGACTGTGTCAGAGGACACCCACTCAACTGGAAGAGCTG 1074  
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
Db 1075 GACAGCTCCCTCTGTTTCTGAGAGTGCACAGGGGAGGAGTCTCTCTCAGTGGGT 1134  
Qy 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
Db 1135 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGAT 1194  
Qy 261 Ala 261  
Db 1195 GCC 1197  
RESULT 10  
AX780857 2788 bp DNA linear PAT 14-JUL-2003  
LOCUS Sequence 3014 from Patent WO03039443.  
DEFINITION AX780857  
ACCESSION AX780857  
VERSION AX780857.1 GI:32697851  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Haerlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
Dugas, M., Ellis, R., Brois, B. and Mergenthaler, S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 3014 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE);  
Ludwig-Maximilian-Universitaet Muenchen (DE);  
PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)  
FEATURES  
source  
Location/Qualifiers  
1..2788  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores: 5.32e-101 Length: 2788  
Pred. No.: 1341.00 Matches: 259  
Score: 99.23% Conservative: 0  
Percent Similarity: 99.23% Mismatches: 2  
Best Local Similarity: 99.11% Indels: 0  
Query Match: 99.11% Gaps: 0  
DB: 6  
US-09-939-853A-75 (1-261) x AX780857 (1-2788)  
Qy 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
Db 387 ATGGGAAGTCTCCCGCAGAGAGAAATCTCTGCAAGCCCAAGCTTGATTCCTCTGTC 446  
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
Db 447 CAAGGCCANGAGCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGSCCCTG 506  
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
Db 507 GGCAGTTTCCCGCAGGTTGGCCCGCCGAGCTGTCTGAGACTCGGGAGCCTTGACC 566

QY	61	IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr	80
DB	567	ATCGTCTCTGAGGATGGAGACTGTGTGACGGTGTCTGAAGTCTCAGGACAGAGATAT	626
QY	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
DB	627	AACATCCCGCCGCTCCACGTCGCCCAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	686
QY	101	ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu	120
DB	687	AGGAGAAAGCAGAGGAACCTCTGTGTACTCTGGGAACCTCGAGGGGCTTCCTCATC	746
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
DB	747	CGGAGAGCCAGACAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCGCCCTGCA	806
QY	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
DB	807	TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACAAATGGCTGGCTGATC	866
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
DB	867	TCACCGCGCTCACTCCCTCTACTCAGGCCCTGGTGACCACTTACTCTGAGCTGGCG	926
QY	181	AspAspIleCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
DB	927	GATGACATCTCTGCTACTCAAGAGGCCCTGTCTCTCAGAGGGCTGGCCGCTCCCT	986
QY	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
DB	987	GGCAGGATATACCCCTTACCTGTACTGTGAGAGGACCCACTCAACTGGAAAGAGCTG	1046
QY	221	AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly	240
DB	1047	GACAGCTCCCTCTCTTTCTTCTGAGCTGCCACAGGGAGGAGTCTCTCTCAGTGAGGT	1106
QY	241	LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp	260
DB	1107	CTCCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGACAGGGCTGTCTCTTTGGATGAT	1166
QY	261	Ala 261	
DB	1167	GCC 1169	
RESULT 11			
AX511155			
LOCUS	AX511155	737 bp	DNA linear PAT 27-SEP-2002
DEFINITION	Sequence 6 from Patent WO0242452.		
ACCESSION	AX511155		
VERSION	AX511155.1	GI:23392047	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	McGlade, J.C. and Loreto, M.P.		
TITLE	Adapter gene		
JOURNAL	Patent: WO 0242452-A 6 30-MAY-2002;		
FEATURES	The Hospital for Sick Children (CA)		
source	Location/Qualifiers		
	1..737		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	9,426-90	Length:	737
Score:	1196.50	Matches:	240
Percent Similarity:	92.0%	Conservative:	2
Best Local Similarity:	91.25%	Mismatches:	2

Query Match:	88.43%	Indels:	19
DB:	6	Gaps:	1
US-09-939-853A-75 (1-261) x AX511155 (1-737)			
QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal	20
DB	1	ATGGGAAGTCTCCCGCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAGTTCTCTGTGC	60
QY	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
DB	61	CAAGGCCAGGGACCTGTGTACCATGGAAGCAGAGAAAGCAAGCCACAGCCCTGGCCCTG	120
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
DB	121	GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTCTGCTGAGACTCGGGAGCAATTGACC	180
QY	61	IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr	80
DB	181	ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGAAAGTCTCAGGCAGAGAGATAT	240
QY	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
DB	241	AACATCCCGCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	300
QY	101	ArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu	120
DB	301	AGGGAGAAAGCAGAGGAACCTGTGTGTGTACCTGGGAACCTCGAGGGGCTTCCTCATC	360
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
DB	361	CGGGAGAGCCAGACCCAGAGAGGCTTACTCTCTGTCTGCTCAGTCCGCTCAGCCGCTGCA	420
QY	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
DB	421	TCTGGACCGGATACAGACACTACAGGATCCACTGCTTGAATGATGGTGGCTGTACATC	480
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
DB	481	TACCGCGCTCACCTTCCCTCTACTCCAGGCCCTGGTGGACCATTC-----	528
QY	181	AspAspIleCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr	200
DB	529	-----TCTGAGGGCTGGCCCGCTCCC	549
QY	200	OGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe	220
DB	550	TGGCAGAGGATATACCCCTACCTGTGACTGTGGGAGGACACCACCTCACTGGAAAGAGCT	609
QY	220	uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluG	240
DB	610	GGACAGCTCCCTCTGTCTTCTGAACTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGG	669
QY	240	YLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA	260
DB	670	TCTCCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGAGCGAGGCTGTCTCTTTGGATG	729
QY	260	spAla 261	
DB	730	ATGCC 734	
RESULT 12			
AF290986			
LOCUS	AF290986	737 bp	mRNA linear PRI 21-JAN-2003
DEFINITION	Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,		
	complete cds; alternatively spliced.		
ACCESSION	AF290986		
VERSION	AF290986.1	GI:17351922	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

REFERENCE 1 (bases 1 to 737)  
 Loreto, M.P. and Mcglade, C.J.  
 Cloning and characterization of human Src-like adaptor protein 2  
 and a novel splice isoform, SLAP-2-v  
 Oncogene 22 (2), 266-273 (2003)  
 MEDLINE 22415750  
 PUBMED 12527895  
 REFERENCE 2 (bases 1 to 737)  
 Loreto, M.P. and Mcglade, C.J.  
 Direct Submission  
 TITLE Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for  
 Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1. 737  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /tissue\_type="thymus"  
 1. 633  
 /note="SLAP-2-v"  
 /codon\_start=1  
 /product="Src-like adaptor protein-2 splice isoform"  
 /protein\_id="AAL38198.1"  
 /db\_xref="GI:17351923"  
 /translation="MGSLEPRKRSKSLPSPSSVSGQGPVTMEAKSKATVALGSPFP  
 AGPAEISLRIGELTIVSDGDWTVLSEVSGEYNIPSHVAKVSHGWLVEGLSRE  
 KAEELLILPQNGAFILIRISQTRGYSLSVRLSPASWDRIHRIHCLDNGWLYI  
 SPRLTFPSLQALVDHYSEGWPAPQWGTPTCDCAEDTTQLERAGQLPPVF"

Alignment Scores:  
 Pred. No.: 9, 42e-90 Length: 737  
 Score: 1196.50 Matches: 240  
 Percent Similarity: 92.02% Conservative: 2  
 Best Local Similarity: 91.25% Mismatches: 2  
 Query Match: 88.43% Indels: 19  
 DB: 9 Gaps: 1

US-09-939-853a-75 (1-261) x AF290986 (1-737)

QY 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 60  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 61 CAAGGCCAGGGACCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGlyGluProLeuThr 60  
 Db 121 GCACGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 181 ATCGTCTCTGAGGATGGAGACTGGTGAGCGGTGCTGTAAGTCTCAGGAGAGAGAT 240  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
 Db 241 AACATCCCAGAGTCCAGCTGGCCAAAGTCTCCCATGGTGGTGTATGAGGGCTGAGC 300  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 Db 301 AGGGAGAAAGCAGAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 361 CGGGAGAGCCAGACAGAGAGAGCTTACTTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 420  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 Db 421 TCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGTGACATGGTGGCTGTGATCATC 480

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 481 TCACGGCCCTCACTCCCTCCCTCACTCCAGCCCTCGTGGACCATTTAC----- 528  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200  
 Db 529 -----TCTGAGGGCTGGCCCGCTCCC 549  
 QY 200 cglLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220  
 Db 550 TGGCAGAGATATACCCCTACTCTGTGACTGTGGAGAGACACACTCACTCACTGGAAGAGCT 609  
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGl 240  
 Db 610 GGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTCTCAGTGAGGG 669  
 QY 240 vLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA 260  
 Db 670 TCTCCGGAGTCCCTCACTTCTACATCAGCTGATGAGCGAGCTCTCTCTTTGGATG 729  
 QY 260 spAla 261  
 Db 730 ATGCC 734  
 RESULT 13  
 AX511151  
 LOCUS  
 DEFINITION Sequence 2 from Patent WO0242452.  
 ACCESSION AX511151  
 VERSION AX511151.1 GI:23392045  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS Mcglade, J.C. and Loreto, M.P.  
 TITLE Adapter gene  
 JOURNAL Patent: WO 0242452-A 2 30-MAY-2002;  
 The Hospital for Sick Children (CA)  
 FEATURES  
 source  
 1. 777  
 /organism="Mus musculus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:10090"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 7, 96e-76 Length: 777  
 Score: 1028.00 Matches: 208  
 Percent Similarity: 85.50% Conservative: 16  
 Best Local Similarity: 79.39% Mismatches: 34  
 Query Match: 75.98% Indels: 4  
 DB: 6 Gaps: 3

US-09-939-853a-75 (1-261) x AX511151 (1-777)

QY 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 57  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 58 CCAGACCCAGAACCCGTGTCCATGCGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 117  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 118 GGCAGTTTCCAGCAGGTGAAACAGGCGCAGACTATCTCTGAGACTCGGGAGCGCTGACC 177  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 178 ATCATCTCTGAGGATGGAGATTGGTGACAGTCCAGTCCAGTCCAGTCCAGGAGAGATAC 237

```

QY      51  AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTrpGluGlyLeuSer 100
      Db      238  CACATGCCAGTGTGTATGTGCTAAAGTCCCCACGGGTGGCTGTACGAGGGGCTGAGC 297
QY      101  ArgGluLysAlaGluGluLeuLeuLeuProGlyLysAsnProGlyGlyAlaPheLeuLe 120
      Db      238  CGGAGAAAGCGGAGAACTACTCTCTGTACTGGAAACCCCGAGGGGCTTCTCTCATC 357
QY      121  ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
      Db      358  CGGAGAGCCAGACACGAGAGAGGCTGTATTCCTGTCCGTCCGACTCAGCGGCCCTGCA 417
QY      141  SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
      Db      418  TCTTGGGACCGATCAGACACTACAGGATCAGCGTCTTGCAATGGCTGGCTGTATCATC 477
QY      161  SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
      Db      478  TCACCTCGCCTCACCTTCCCTCACTCCACGCTTGGTGGAGCATTACTCTGAGCTAGCA 537
QY      181  AspAspIleCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
      Db      538  GATGGCATCTGTCTGCCCTCAGGAGCGGTGTCTGCGAGAGCTTGGGCCACTACCT 597
QY      201  GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
      Db      598  GCGAAGATACACCTCCACCTGTGACTGTGCCAATCATCATCTAAATTGGAAAAGCTG 657
QY      221  AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuSerGlu 239
      Db      658  GACCCGAGCCTCTCTTTCTGGAACACCTCGGAGTGGGGAGGCATCTCTCTCAGTGAG 717
QY      240  GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
      Db      718  GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCCTGGCTGAGGAC-----CCCTTGGAT 771
QY      260  AspAla 261
      Db      772  GATGCT 777

RESULT 14
AX511150
LOCUS      AX511150      1348 bp      DNA      linear      PAT 27-SEP-2002
DEFINITION Sequence 1 from Patent WO0242452.
ACCESSION AX511150
VERSION    AX511150.1  GI:23392044
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1
AUTHORS    McGlade,J.C. and Loreto,M.P.
TITLE      Adapter gene
JOURNAL    Patent: WO 0242452-A 1 30-MAY-2002;
           The Hospital for Sick Children (CA)
FEATURES   Location/Qualifiers
            source
              1..1348
              /organism="Mus musculus"
              /mol_type="unassigned DNA"
              /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.:      1,51e-75      Length:      1348
Score:          1028.00      Matches:      208
Percent Similarity: 85.50%      Conservative: 16
Best Local Similarity: 79.39%      Mismatches:  34
Query Match:     75.98%      Indels:      4
DB:              6              Gaps:        3

US-09-939-853A-75 (1-261) x AX511150 (1-1348)

```

```

QY      1  MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal 20
      Db      282  ATGGGAAGTTTGTCCAGCAGAGGGAAACCC---TCCAGCCCCAGCCCCAGCTCCTCTGGT 338
QY      21  GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
      Db      339  CCAGACCAGAACCCGTGTCCATGCAACCAAGAAACACACAGGTTCACAGCTGTGGCCCTG 398
QY      41  GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
      Db      399  GGCAGTTTCCCGACAGGTGACAGGCGCACACTATCTCTGAGACTCGGGGAGCGCTGACC 458
QY      61  IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80
      Db      459  ATCATCTCTGAGATGAGATTGGTGGACAGTCCAGTCCGAAAGTCTCAGCAGAGAGTAC 518
QY      81  AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
      Db      519  CACATGCCAGTGTGTATGTGGCTAAAGTCGCCACCGGTGGCTGTACGAGGCGCTGAGC 578
QY      101  ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120
      Db      579  CGGAGAAAGCGGAGGACTACTCTCTGTACTGGAAACCCCGAGGGGCTTCTCTCATC 638
QY      121  ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
      Db      639  CGGAGAGCCAGACACGAGAGAGGCTGTATTCCTGTCCGTCCGACTCAGCGGCCCTGCA 698
QY      141  SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
      Db      699  TCTTGGGACCGATCAGACACTACAGGATCAGCGTCTTGCAATGGCTGGCTGTATCATC 758
QY      161  SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
      Db      759  TCACCTCGCCTCACCTTCCCTCACTCCACGCTTGGTGGAGCATTACTCTGAGCTAGCA 818
QY      181  AspAspIleCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
      Db      819  GATGGCATCTGTCTGCCCTCAGGAGCGGTGTCTCTCAGAAAGCTTGGGCCACTACCT 878
QY      201  GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
      Db      879  GCGAAGATACACCTCCACCTGTGACTGTGCCAATCATCATCTAAATTGGAAAAGCTG 938
QY      221  AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuSerGlu 239
      Db      939  GACCCGAGCCTCTCTTTCTGGAAGCACCTCGAGTGGGGAGGCATCTCTCTCAGTGAG 998
QY      240  GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
      Db      999  GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCCTGGCTGAGGAC-----CCCTTGGAT 1052
QY      260  AspAla 261
      Db      1053  GATGCT 1058

RESULT 15
AF287467
LOCUS      AF287467      1348 bp      mRNA      linear      ROD 03-JUN-2002
DEFINITION Mus musculus Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF287467
VERSION    AF287467.1  GI:17351918
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1
AUTHORS    Loreto,M.P., Berry,D.M. and McGlade,C.J.
TITLE      Functional cooperation between c-Cbl and Src-like adaptor protein 2
           in the negative regulation of T-cell receptor signaling
JOURNAL    Mol Cell. Biol. 22 (12), 4241-4255 (2002)
MEDLINE    22022020

```

PUBMED 12024036  
REFERENCE 2 (bases 1 to 1348)  
AUTHORS Loreto, M.P. and McGLade, C.J.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada  
FEATURES  
source Location/Qualifiers  
1..1348  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/dev\_stage="day 15 embryo"  
282..1061  
/note="SLAP-2"  
/codon\_start=1  
/product="Src-like adaptor protein-2"  
/protein\_id="AAL38196.1"  
/db\_xref="GI:17351919"  
/translation="MGSLSRGTSSPSSSGDPEPVSMPQERHKVTAVALGSPFA  
GEQARLSLRIGEPITIIISERQDMWTQSEVSGREVHMPVSYYAKVAHGLYGLSREK  
AEELLIPGNPGNAPLIRSQTRGCYSLVRLSRPASWDRIHRYRIQLDNGWLYIS  
PRITFPFLHALVEHYSLEADGICCPUREPCVLOKGLPFGKOTPPPTVTTSLNWKK  
LDRSLDFLEAPSGEASLLSEGLRESLSYSIAEDPLDDA"  
381..551  
/note="Region: SH3 domain"  
552..824  
/note="Region: SH2 domain"  
1301..1306  
polyA\_signal  
ORIGIN  
Alignment Scores:  
Pred. No.: 1348  
Score: 1028.00  
Matches: 1348  
Percent Similarity: 85.50%  
Conservative: 16  
Best Local Similarity: 79.39%  
Mismatch: 34  
Query Match: 75.98%  
Indels: 4  
Gaps: 3  
DB: 10  
US-09-939-853A-75 (1-261) x AF287467 (1-1348)  
QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal 20  
DB 282 ATGGAGAGTTGTCAGCAGAGGAGAAAC---TCCAGCCCCCAGCCAGCTCTCTGGT 338  
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
DB 339 CCAGACCAGGAACCGCTGCCATGCACACCAAGAACACACAGGTCACAGCTGTGGCCCTG 398  
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
DB 399 GGCAGTITCCAGCAGGAGGAGACAGCCAGACTATCTGTGACTCCGGGAGCCGCTGACC 458  
QY 61 IleValSerGluAspGlyAspTyrTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
DB 459 ATCATCTCTAGGATGGAGATTGTGGACAGTCCAGTCCGAAAGTCTCAGGCAGAGAGTAC 518  
QY 81 AsnIleProSerValHisValGlyValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
DB 519 CACATGCCACGTGTGTATGTGGCTAAAGTCGCCACGGGTGGCTGTACGAGGGGCTGAGC 578  
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120  
DB 579 CGGAGAAAGCCGAGGAATACTCTGTGTACTGGGAACCCCGGAGGGGCTTCTCTATC 638  
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
DB 639 CGGGAGACCCAGACAGAGAGGCTGTATTCCCTGTCCGCTCCGACTCAGCCGCCCTGCA 698  
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
DB 699 TCTTGGGACCGGATCAGACACTACAGGATACAGCGCTTGTGACATGGCTGGCTGATCATC 758  
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180

DB 759 TCACCTCGCCTCACCTTCCCTCCTCCACGCTTGGTGGAGCATTACTCTGAGCTAGCA 818  
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
DB 819 GATGGCATCTGCTGTCCCTCAGGAGCCGCTGTCTCTGCAGAGCTTGGGCCACTACCT 878  
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
DB 879 GGCRAAGATACACCTCCACCTGTGACTGTGCCAACATCATCATCAATTTGAAAAGCTG 938  
QY 221 AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuSerGlu 239  
DB 939 GACCGCAGCTTCTTCTGAGAGCCTCGGAGTGGGAGGCACTCTCTGCTCAGTGAG 998  
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259  
DB 999 GGGCTCCGAGAGTCCCTCAGTTCCTATCATCAGCCTGGCTGAGGAC-----CCCTTGGAT 1052  
QY 260 AspAla 261  
DB 1053 GATGCT 1058

Search completed: February 20, 2004, 10:48:33  
Job time : 3102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:33:07 ; Search time 368 Seconds  
(without alignments)  
3012.988 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSILPSRRKSLPSPSLSSV.....RESLSPVISLNDRAVSLDDA 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO.spool p/US09939853/runat 19022004.145338 24420/app query.fasta\_1.455  
-DB=N Geneseq 29Jan04 -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09939853.ecgn 1.1708 @runat 19022004.145338 24420 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DGPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1980s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1353	100.0	1183	6	ABK61465 Human CDN
2	1347	99.6	786	6	AAL44089 Human ORF
3	1347	99.6	786	6	ABQ74343 Human Src
4	1347	99.6	2567	6	AD43980 Human Src
5	1269	93.8	837	3	AAC77202 Human ORF
6	1206.5	89.2	1413	6	ABQ99374 Human cod
7	1196.5	88.4	737	6	AAL44090 Mouse MAR
8	1128.5	83.4	2049	5	AAS74750 DNA encod

9	1028	76.0	1348	6	AAL44087 Mouse mod
10	826	61.0	763	6	ABQ98670 Human ORF
11	731	54.0	603	5	AAS74748 DNA encod
12	586	43.3	875	6	ABQ99151 Human ORF
13	488	36.1	3756	7	ABX62975 Human act
14	487	36.0	2109	4	AAS02049 DNA encod
15	487	36.0	2665	6	ABL65189 Lung can
16	487	36.0	2665	6	ABX83738 Human CDN
17	487	36.0	2665	7	ACC81091 Human ORF
18	452.5	33.4	444	6	ABQ98669 Human Src
19	370.5	27.4	2298	6	ABK83935 Human CDN
20	370.5	27.4	2298	9	ADD19015 Human dis
21	360.5	26.6	1926	6	ABK83940 Human CDN
22	360.5	26.6	2015	6	ABL66673 Lung can
23	360.5	26.6	2015	6	ABK83939 Human CDN
24	360.5	26.6	2015	9	AD62155 Human hae
25	360.5	26.6	2341	9	ABE53793 Human pro
26	358.5	26.5	1911	6	ABK63704 Rat sequ
27	358.5	26.5	1911	9	ADB58205 Toxicity-
28	358.5	26.5	1911	9	ADB52725 Primary r
29	348	25.7	1530	8	ADA02968 Human LCK
30	348	25.7	1530	9	ADB72706 Human LCK
31	348	25.7	1530	9	ADC85448 Human LCK
32	348	25.7	2032	3	AAX46491 PKA subst
33	348	25.7	2032	7	ACC72850 Human can
34	348	25.7	2032	8	ADA02967 Human LCK
35	348	25.7	2032	9	ADB72705 Human LCK
36	348	25.7	2032	9	ADC85447 Human LCK
37	348	25.7	2032	9	ADE40448 Human pro
38	340	25.1	2034	6	ABG65202 CDNA enco
39	340	25.1	2129	7	ACA56854 Human sig
40	340	25.1	2129	7	ACC81082 Human t-1
41	340	25.1	2129	7	ACA64888 Human p56
42	338.5	25.0	1530	8	ADA02965 Mouse Lck
43	338.5	25.0	1530	9	ADB72703 Mouse Lck
44	338.5	25.0	1530	9	ADC85445 Mouse Lck
45	338.5	25.0	2100	8	ADA02964 Mouse Lck

ALIGNMENTS

RESULT 1

ABK61465  
ID ABK61465 standard; CDNA; 1183 BP.

XX AC ABK61465;

XX DT 18-JUN-2002 (first entry)

XX DE Human cDNA encoding protein NOV13.

XX KW Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;

XX KW cell signal processing disorder; metabolic pathway modulation disorder;

XX KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;

XX KW uterus cancer; immune response; graft-versus-host disease;

XX KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;

XX KW hypertension; congenital heart defects; multiple sclerosis; inflammation;

XX KW Albright hereditary osteodystrophy.

XX OS Homo sapiens.

XX PN WO200216599-A2.

XX PD 28-FEB-2002.

XX PF 27-AUG-2001; 2001WO-US026510.

XX PR 25-AUG-2000; 2000US-0228191P.

XX PR 08-FEB-2001; 2001US-0267300P.

XX PR 20-FEB-2001; 2001US-0269961P.

XX PR 20-MAR-2001; 2001US-0277337P.

XX PA (CURA-) CURAGEN CORP.

(CORT-) COR THERAPEUTICS INC.

PA  
XX  
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;  
FI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;  
XX  
DR WPI; 2002-280937/32.  
DR P-PSDB; AAU91308.

XX  
XX New polypeptides for treating or preventing a disorder associated with  
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.  
PT  
XX

PS Claim 1; Page 98; 263pp; English.

XX  
XX The invention relates to an isolated polypeptide (NOVX) a mature form of  
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide  
CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
CC 15a, 15b, 16a, and 16b, the NOVX polypeptide, nucleic acid encoding it,  
CC and antibody against it, are useful for treating or preventing (e.g. by  
CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
CC atherosclerosis, a disorder related to cell signal processing and  
CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
CC and nucleic acids are also useful for determining the presence of  
CC predilection to the diseases. The NOVX nucleic acid and polypeptide are  
CC especially useful in therapeutic or prophylactic applications for  
CC disorders associated with aberrant NOVX expression or activity, e.g.  
CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
CC cancer), immune response, graft-versus-host disease, acquired  
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
CC congenital heart defects, multiple sclerosis, inflammation or Albrit  
CC hereditary osteodystrophy and many other diseases listed in the  
CC specification. The DNA encoding the protein is useful in gene therapy for  
CC treating the conditions. This is also useful in detection assays,  
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
CC for developing a powerful assay system for functional analysis of various  
CC human disorders, as well as in diagnostic applications. The present  
CC sequence encodes a NOVX protein

XX  
SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,41e-109 Length: 1183  
Score: 1353.00 Matches: 261  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x ABK61465 (1-1183)

QY 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
DB 398 ATGGAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 457  
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
DB 458 CAGGCCAGGACCTGTGACCATGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 517  
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGlyGluProLeuThr 60  
DB 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCTGAGACTCGGGAGCCATTGACC 577  
QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
DB 578 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGTAAGTCTCAGGAGAGAGAT 637  
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
DB 638 AACATCCCGCAGCTCCAGCTGGGCAAGTCTCCCATGGTGGCTGTATGAGGCGCTGAGC 697  
QY 101 ArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLe 120  
DB 698 AGGGAGAAAGCAGAGAACTGCTGTGTGTACCTGGGAACCCCTGGAGGGGCTTCTCTCATC 757

QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
DB 758 CGGGAGAGCCAGACCCAGAGAGAGCTCTTACTCTCTGTCTGCTCCGCTCAGCGCCCTGCA 817  
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspGlyTrpLeuTyrIle 160  
DB 818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCATATGGCTGGCTGTATC 877  
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
DB 878 TCACCGCGCTCACCTCCCTCCTCCTCAGGCTGCTGCTGACCATTTACTCTGAGCTGGCG 937  
QY 181 AspAspIleCysCysLeuLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
DB 938 GATGACATCTGCTCTCTCTCAGGAGCTGTCTCTGAGAGGCTGGCCCTCCCT 997  
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
DB 998 GGCAGGATATACCTTACCTGTGACTGTGAGAGGACACCCTCAACTGGAAGAGCTG 1057  
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuSerGluGly 240  
DB 1058 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAGGT 1117  
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260  
DB 1118 CTCGGGAGTCCCTCAGCTTCTATCAGCTGATGACGAGGCTGTCTCTTGGATGAT 1177  
QY 261 Ala 261  
DB 1178 GCC 1180  
RESULT 2  
AAL44089  
ID AAL44089 standard; cDNA; 786 BP.  
XX  
AC AAL44089;  
XX  
DT 03-OCT-2002 (first entry)  
XX  
DE Human modulator of antigen receptor signalling protein coding sequence.  
XX  
KW Human; gene; ss; gene therapy; modulator of antigen receptor signalling;  
KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP; disorder;  
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
KW immunosuppression; myeloproliferative disorder; breast cancer.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..786  
FT /\*tag= a  
FT /product= "Human MARS protein"  
XX  
PN WO200242452-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 26-NOV-2001; 2001WO-CA001662.  
XX  
PR 27-NOV-2000; 2000CA-02324663.  
XX  
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.  
XX  
PI Mcglade JC, Loreto MP;  
XX  
FI WPI; 2002-566564/60.  
DR P-PSDB; AAO15457.  
XX  
PT New isolated modulator of antigen receptor signaling protein or its  
PT fragment, useful for treating malignant disorders such as myeloid  
PT malignancies, autoimmune disorders and myeloproliferative disorders.  
XX

PS Claim 12; Page 75; 110pp; English.

XX The invention comprises the amino acid and coding sequences of modulator  
CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
CC putative tumour suppressor gene and exhibits structural and sequence  
CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and  
CC protein sequences of the invention are useful for the treatment of  
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
CC disorders, immunosuppression, myeloproliferative disorders and  
CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
XX breast cancer). The present cDNA sequence encodes a human MARS protein

SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.79e-109 Length: 786  
Score: 1347.00 Matches: 260  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.56% Indels: 0  
DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x AAL44089 (1-786)

Qy 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
Db 1 ATGGGAAGTCTGCCAGCAGAGAATAATCTCTGCAAGCCCAAGCTTGTCTCTGTC 60  
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40  
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCAGCGCTGGCCCTG 120  
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGlyGluProLeuThr 60  
Db 121 GCGAGTTTCCCGGAGGTGGCCGGCCAGCTGTCTGCTGAGACTCGGGAGCCATTGACC 180  
Qy 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
Db 181 ATCTCTCTGAGGATGGAGACTGTGGACGGTGTCTGTGAAGTCTCAGGACGAGAGATAT 240  
Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
Db 241 AACATCCCGCCAGTCCAGCTGGCCAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGC 300  
Qy 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLe 120  
Db 301 AGGAGAGAAGCAGAGAACTGCTGTGTGTACTGGAAACCCCTGGAGGGGGCTTCTCTCATC 360  
Qy 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
Db 361 CGGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTGTCAGTCCGCCCTCAGCCGCCGCA 420  
Qy 141 SerTipAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160  
Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCTTGACAATGGCTGGCTGTACATC 480  
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
Db 481 TCACCGCCCTCACCTTCCCTCTCATCCAGGCCCTGGTGGACCAATTACTCTGAGCTGGCG 540  
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
Db 541 GATGACATCTGCTGCTACTCAAGGAGCCCTGTGTCTGCGAGAGGCTGGCCCGCTCCCT 600  
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
Db 601 GCGAAGGATATACCCCTACCTGTGACTGTGCGAGAGGACACCACCTCACTGGAAGAGCTG 660  
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
Db 661 GACAGCTCCCTCTCTCTTCTGAGACTGCCACAGGGAGAGTCTCTCTCAGTGAGGCT 720  
Qy 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260

Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGACGAGGCTGTCTCTTGGATGAT 780  
Qy 261 Ala 261  
Db 781 GCC 783  
RESULT 3  
ABQ74343  
ID ABQ74343 standard; cDNA; 786 BP.  
XX AC ABQ74343;  
XX 15-OCT-2002 (first entry)  
XX Human Src-like inhibitory molecule (SLIM) encoding cDNA.  
DE Human; Src-like inhibitory molecule; SLIM; Src-like adaptor protein;  
XX SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;  
KW modulator; lymphocyte; Cbl; Gene therapy; immunodeficiency disorder;  
KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;  
KW Gene; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX CDS 1..786  
XX /\*tag= a "SLIM"  
XX /product= "SLIM"  
XX /note= "Src-like inhibitory molecule"  
PN WO200255707-A2.  
XX 18-JUL-2002.  
XX 10-JAN-2002; 2002WO-US000718.  
XX 10-JAN-2001; 2001US-0260953P.  
XX (RIGE-) RIGEL PHARM INC.  
XX Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
XX Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
XX WPI; 2002-575432/61.  
XX P-PSDB; ABP52187.  
XX New src-like inhibitory molecule protein, useful for treating  
XX immunodeficiency disorders and inflammatory disorders, comprises N-  
XX terminal myristylation sequence, SH2 domain and/or SH3 domain.  
XX Claim 3; Fig 2A; 91pp; English.  
XX The present sequence encodes the human Src-like inhibitory molecule  
XX (SLIM) protein (I). The present invention describes a SLIM protein  
XX comprising an N-terminal myristylation sequence, an N-terminal SH2  
XX domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising  
XX an N-terminal myristylation sequence and an N-terminal SH2 domain which  
XX is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and  
XX anti-HIV activities, and can be used as a modulator of lymphocyte  
XX activation, and of ubiquitination of a Cbl target protein, and in gene  
XX therapy. (I) is useful for screening a bioactive agent capable of binding  
XX to SLIM. (I) is also useful for screening a bioactive agent capable of  
XX modulating SLIM binding. (I) or its fragments is useful in the study or  
XX in the treatment of conditions which involves this function or  
XX dysregulation of SLIM protein activity, i.e. to diagnose, treat or  
XX prevent SLIM associated disorders. (I) or the polynucleotide encoding it  
XX (II) is useful for modulating leukocyte and/or platelet activation, for  
XX modulating antigen receptor-induced signalling and activation in  
XX leukocyte and/or platelets and for modulating antigen receptor-induced  
XX signalling and activation in lymphocytes and/or mast cells. (I) or (II)  
XX is also useful for modulating the basal activity of lymphocytes. (I) or



CC (II) is useful in the treatment of immunodeficiency disorders, such as  
 CC acquired immunodeficiency syndrome (AIDS), for the prevention and  
 CC treatment of acute inflammatory disorders, chronic inflammatory  
 CC disorders, autoimmune disorder and transplant rejection  
 XX  
 SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4, 79e-109 Length: 786  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 99.62% Conservatives: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.56% Indels: 0  
 DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x ABQ74343 (1-786)

Qy 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAATAATCTCTGCAAGCCCAAGCTTGATCTCTCTGTC 60

Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 61 CAAGCCAGGAGCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCAGCCGTGGCCCTG 120

Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 121 GGCAGTTTCCCGCAGGTGGCCGCGAGCTGTCTGCTGAGACTCGGGGAGCCATTGACC 180

Qy 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 181 ATCTCTCTGTAGGATGGAGACTGTGTGACCGGTGTCTGAAGTCTCAGGCACAGAGTAT 240

Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 Db 241 AACATCCCGCAGCTCCAGTGGCCGCAAGTCTCCATGGTGGCTGTATGAGGCCCTGAGC 300

Qy 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 Db 301 AGGAGAAAGCAGAGGAAGTCTCTGTGTACCTGGGAACCTGGAGGGGCTTCTCCATC 360

Qy 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 361 CGGAGAGCCAGACACAGAGAGGCTCTTACTCTCTGTCTGCTCAGTCCGCTCAGCCGCTGCA 420

Qy 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 Db 421 TCCTGGGACCGGATCAGACATACAGGATCCACTGCTTGCACATGGCTGGCTGTATC 480

Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 481 TCACCGGCTTCACTTCCCTCTACTCCAGCCCTGTGTGACCACTTACTGTAGCTGGCG 540

Qy 181 AspAspIleCysLysLeuLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 Db 541 GATGACATCTGTGCTACTCAAGGAGCCCTGTCTCTGACAGAGGCTGGCCGCTCCCT 600

Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 Db 601 GCAAGGATATACCCCTACCTGTGACTGTGACAGAGGACACCACTCAACTGAAAGAGCTG 660

Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
 Db 661 GACAGCTCCCTCTCTTCTTCTGAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGCT 720

Qy 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
 Db 721 CTCGGGAGTCCCTCAGCTTCTATCATGACGCTGAATGACAGGCTGTCTCTTGGATGAT 780

Qy 261 Ala 261  
 Db 781 GCC 783

RESULT 4  
 AAD43980  
 ID AAD43980 standard; cDNA; 2567 BP.  
 XX  
 AC AAD43980;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 DE Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.  
 XX  
 KW Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 415..1200  
 FT /\*tag= a  
 FT /product= "Human SLAP-2"  
 XX  
 PN WO200242457-A1.  
 XX  
 PD 30-MAY-2002.  
 XX  
 PF 20-NOV-2001; 2001WO-US043367.  
 XX  
 PR 22-NOV-2000; 2000US-0252545P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;  
 XX  
 DR WPI: 2002-463632/49.  
 XX  
 DR P-PSDB; AAE26357.  
 XX  
 PT Novel substantially purified human SH2/SH3-domain-containing adapter  
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic  
 PT intervention in immunological and inflammatory disorders and cancer.  
 XX  
 PS Claim 2; Fig 1; 85pp; English.  
 XX  
 CC The invention relates to a substantially purified human SH2/SH3-domain-  
 CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-  
 CC 2). The invention is useful for treating an immune disorder involving  
 CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is  
 CC useful for screening for antagonists or inhibitors of the interaction of  
 CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or  
 CC preventing diseases or disorders associated with aberrant or uncontrolled  
 CC cellular signal transduction, for determining those cellular signalling  
 CC molecules which associate with hSLAP-2 and which provide critical signals  
 CC for cell activation, and as effectors in methods to affect T- cell  
 CC activation. The invention is useful in screening assays to identify and  
 CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for  
 CC potential use to treat autoimmune diseases which may be caused by  
 CC hyperactivated B cells, as well as to treat diseases which may be caused  
 CC by hyperactivated T cells, in addition to other immune system related  
 CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,  
 CC inflammation disorders, diseases and conditions, rheumatoid arthritis,  
 CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's  
 CC and ulcerative colitis), allergies, particularly those involving  
 CC hyperactivity of B-cells and T- cells, or other immune cells, such as  
 CC mast cells or eosinophils, autoimmune diseases such as systemic lupus  
 CC erythematosus and multiple sclerosis, pulmonary diseases including  
 CC asthma, acute respiratory distress syndrome, and chronic obstructive  
 CC pulmonary disorder, tissue/ organ rejection and cancer. The invention is  
 CC useful in gene therapy. The present sequence is human SLAP-2 cDNA  
 XX  
 SQ Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Alignment Scores:		2.19e-108	Length:	2567
Pred. No.:	Score:	1347.00	Matches:	260
Percent Similarity:	Score:	99.62%	Conservative:	0
Best Local Similarity:	Score:	99.62%	Mismatches:	1
Query Match:	Score:	99.56%	Indels:	0
DB:	Score:	6	Gaps:	0
US-09-939-853A-75 (1-261) x AAD43980 (1-2567)				
QY	1	MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal	20	
DB	415	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTCAGTTCCTCTGTC	474	
QY	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40	
DB	475	CAAGGCCAGGACCTGTGACCATGTAACACAGAGAGAACAGGCCACAGCCGCTGGCCCTG	534	
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGlyGluProLeuThr	60	
DB	535	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCTGAGACTCGGGAGCCATTGACC	594	
QY	61	IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr	80	
DB	595	ATCGTCTCTGAGATGAGACTGTGTGGACGGTGTGTCTGAAGTCTCAGGCAGAGATAT	654	
QY	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100	
DB	655	AACATCCCAAGCTCCAGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC	714	
QY	101	ArgGlyLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle	120	
DB	715	AGGGAGAAAGCAGAGAACTGTGTGTGTACTTGGGAACCTGGAGGGGCTTCTCTCATC	774	
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140	
DB	775	CGGGAGAGCCAGCAGGAGAGGCTTCTACTCTGTGTCAGTCCGCTCAGCCGCTGCA	834	
QY	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160	
DB	835	TCTTGGGACCGATCAGACACTACAGGATCCACTGCCCTTGACATGGCTGTGTACATC	894	
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180	
DB	895	TCACCGCGCTCACCTTCCCTCACTCAGGCGCCCTGGTGGACCATTAATCTGAGCTGGCG	954	
QY	181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200	
DB	955	GATGACATCTGTGCTACTCAGGAGCCCTGTGCTGCAGAGGCTGGCCCGCTCCCT	1014	
QY	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220	
DB	1015	GGCAAGGATATACCCCTCACTGTGACTGTGCAGAGGACACCACTCAATGGAAAGAGCTG	1074	
QY	221	AspSerSerLeuLeuPheSerGluAlaThrGlyGluGluSerLeuLeuSerGluGly	240	
DB	1075	GACAGCTCCCTCTCTTTTCTGAACTGCCACAGGGAGAGTCTCTCTCAGTGAGGGT	1134	
QY	241	LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp	260	
DB	1135	CTCCGGAGTCCCTCAGCTTCTTACATCAGCCTGAATGACGAGCTGTCTCTTTGGATGAT	1194	
QY	261	Ala 261		
DB	1195	GCC 1197		
RESULT 5				
ID	AAC77202 standard; cDNA; 837 BP.			
XX	AAC77202			
AC	AAC77202;			
XX				

DT	08-FEB-2001	(first entry)
XX	Human ORFX ORF2757 polynucleotide sequence SEQ ID NO:5513.	
DE	Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;	
XX	vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;	
KW	anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;	
KW	immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;	
KW	hypotensive; dermatological; immunosuppressive; antiinflammatory;	
KW	antiviral; antibacterial; antifungal; antirheumatic; antithyroid;	
KW	antianemic; gene therapy; cancer; proliferative disorder; hypertension;	
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;	
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; AIDS;	
KW	cholesterol ester storage; systemic lupus erythematosus; infection;	
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;	
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;	
KW	bone damage; cartilage damage; antiinflammatory disease; coagulation;	
KW	thrombosis; contraceptive; ss.	
OS	Homo sapiens.	
XX	WC200058473-A2.	
XX	05-OCT-2000.	
XX	31-MAR-2000; 2000WO-US008621.	
XX	31-MAR-1999; 99US-0127607P.	
PR	02-APR-1999; 99US-0127636P.	
PR	03-APR-1999; 99US-0127728P.	
PR	30-MAR-2000; 2000US-00540763.	
XX	(CURA-) CURAGEN CORP.	
XX	Shimkets RA, Leach M;	
XX	WPI; 2000-602362/57.	
DR	P-PSDB; AAB42993.	
XX	Novel nucleic acids and peptides derived from open reading frame X,	
PT	useful for treating e.g. cancers, proliferative disorders,	
PT	neurodegenerative disorders and cardiovascular disease.	
XX	Claim 5; Page 4692-4693; 5507pp; English.	
PS	AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,	
CC	which represent the human ORFX open reading frames 1 to 3161. The ORFX	
CC	sequences have activities such as: cytostatic; hepatotropic; vulnary;	
CC	antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;	
CC	anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;	
CC	cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;	
CC	dermatological; immunosuppressive; antiinflammatory; antibacterial;	
CC	antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The	
CC	sequences can be used for determining the presence of or predisposition	
CC	to, or preventing or treating pathological conditions associated with an	
CC	ORFX-associated disorder. The nucleic acids can be used to express ORFX	
CC	proteins in gene therapy vectors. The proteins and nucleic acids may be	
CC	used to treat cancers, proliferative disorders, neurodegenerative	
CC	disorders, osteoarthritis, graft vs host disease, cardiovascular disease,	
CC	diabetes mellitus, hypertension, hypothyroidism, cholesterol ester	
CC	storage, systemic lupus erythematosus, severe combined immunodeficiency	
CC	(SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune	
CC	disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and	
CC	cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to	
CC	enhance coagulation; to inhibit thrombosis; and as a contraceptive	
XX	Sequence 837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;	
Alignment Scores:		
Pred. No.:	3.67e-102	Length: 837
Score:	1269.00	Matches: 244
Percent Similarity:	99.19%	Conservative: 1
Best Local Similarity:	98.79%	Mismatches: 2

cytostatic; haemostatic; virucide; antibacterial; fungicide;  
immunostimulant; cerebroprotective; gene therapy; gene; ss.

Homo sapiens.

WO200259260-A2.

01-AUG-2002.

16-NOV-2001; 2001WO-US042950.

17-NOV-2000; 2000US-00714936.

(HYSE-) HYSEQ INC.

Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

WPI; 2002-590824/63.  
N-PSTDB; ABP64788.

New isolated polynucleotide, useful in research, diagnostic or  
therapeutic methods, e.g. preventing or treating disorders involving  
aberrant protein expression or biological activity.

Claim 1; SEQ ID NO 107; 394pp; English.

The present invention relates to novel human coding sequences (ABQ99268-  
ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in  
therapeutic, diagnostic and research methods. The polynucleotides may be  
used in the field of molecular biology as hybridisation probes, primers  
for PCR, for chromosome and gene mapping, for the recombinant production  
of protein, or in generation of anti-sense DNA or RNA. The  
polynucleotides are useful in diagnostics as expressed sequence tags  
(ESTs) for identifying expressed genes or for physical mapping of the  
human genome. The proteins may be used as molecular weight markers, or as  
nutritional sources or supplements. The proteins may be used to maintain  
and expand cell population in a totipotential or pluripotential state,  
useful for re-engineering damaged or diseased tissues, transplantation,  
manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
polynucleotides and proteins are useful for preventing, treating or  
ameliorating disorders involving aberrant protein expression or  
biological activity, e.g. haematopoietic disorders, central/peripheral  
nervous system diseases, mechanical and traumatic disorders, non-healing  
wounds, immune deficiencies and disorders, infectious diseases caused by  
viral, bacterial or fungal infection, autoimmune disorders, allergic  
reactions and conditions, coagulation disorders, or cancer. The  
polynucleotide sequences of the invention were assembled from ESTs  
isolated mainly by sequencing by hybridisation, and in some cases,  
sequences obtained from one or more public databases. Note: The sequence  
data for this patent did not form part of the printed specification, but  
was obtained in electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pat\_sequences

Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:
Score: 2,21e-96	1413
Percent Similarity: 1206.50	Matches: 240
Best Local Similarity: 91.60%	Conservative: 0
Query Match: 89.17%	Mismatches: 4
DB: 6	Indels: 18
	Gaps: 1

US-09-939-853A-75 (1-261) x ABQ99374 (1-1413)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20

Db 54 ATGGAGGTCTGCCACGACGAGAAGAAATCTGCACCAAGCCAGCTTGAGTTCCTCTC 113

QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40

Db 114 CAAGGCCAGGGACCTGTGACCATTGAAGCAGAGAAGCAAGCGCACCGCTGCCCTG 171

Query Match: 93.79% Indels: 0

DB: 3 Gaps: 0

US-09-939-853A-75 (1-261) x AAC77202 (1-837)

QY 15 SerLeuSerSerValGlnGlnGlyProValThrMetGluAlaGluArgSerLys 34

Db 4 ACCTTGAGTTCTCTGTCCAGGCCAGGACCTGTGACCATGGGAAGCAGAGAACGAG 63

QY 35 AlaThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeuSerLeuArg 54

Db 64 GCCACAGCGCGCGCTGGCGAGTTCCCGCAGGTGGCCGCGGAGCTGTGCTCAGAGA 123

QY 55 LeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrValLeuSerGlu 74

Db 124 CTCGGGGAGCCATTGACCATCTCTCTGAGGTGAGACTGGTGCAGCGTCTCTGAA 183

QY 75 ValSerGlyArgGluTyraSnIleProSerValHisValGlyLysValSerHisGlyTrp 94

Db 184 GTCTCAGGCAGAGAGATAAATCATCCCCAGCGTCCACGTGGCAAAGTCTCCATCGGTGG 243

QY 95 LeuTyrglyLeuSerArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnPro 114

Db 244 CTGTATGAGGGCTGAGCAGGAGGAGAAACAGAGGACCTGCTGTTACCTGGGAACCT 303

QY 115 GlyGlyAlaPheIleuLeuArgGluSerGlnThrArgArgGlySerTyrrSerLeuSerVal 134

Db 304 GGAGGGGCTTCTCATCGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTCTCAGTC 363

QY 135 ArgLeuSerArgProAlaSerTrpAspArgIleArgHISrtyrArgIleHisCysLeuAsp 154

Db 364 CGCCTCAGCGCGCTGCTGCTGGACCGGATCAGACACTACAGATCCACTGCTTGAC 423

QY 155 AsnGlyTrpLeuTyrlleSerProArgLeuthrPheProSerLeuGlnAlaLeuValasp 174

Db 424 AATGGCTGGCTGTATCATCCCGCGCTCACCTTCCCCTCACCCAGGCGCTGGTGAC 483

QY 175 HisTyrrSerGluLeuAlaAspAspileCysCysLeuLeuLysGluProCysValleuGln 194

Db 484 CATTACTCTGAGCTGGCGGATCACAATCTGCTGCTTACTCAAGGACCGCTGTGCTGCAG 543

QY 195 ArgAlaGlyProLeuProGlyLysAspileProLeuProValThrValGlnArgThrPro 214

Db 544 AGGGCTGGCGCGCTGCTGGCAGGATATACCCCTACTGTGCTGCTGACGACACCA 603

QY 215 LeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGlu 234

Db 604 CTCACCTGAAAAGAGCTGACAGAGCTCCCTCTGTTTCTTGAAGCTGCCACAGGGGAGGAG 663

QY 235 SerLeuLeuSerGluGlyLeuArgGluSerPheTyrrIleSerLeuAsnAspGlu 254

Db 664 TCCTTCTCAGTGAGGGTCTCGGAGATCCCTCAGCTTCTATCAGCTGCTGAATGACGAG 723

QY 255 AlavalSerLeuAspAla 261

Db 724 GCTGTCTCTTTGATGATGCC 744

RESULT 6

ABQ99374

ID ABQ99374 standard; cDNA; 1413 BP.

XX AC ABQ99374;

XX DB 25-FEB-2003 (first entry)

DE Human coding sequence SEQ ID 107.

XX Human; expressed sequence tag; EST; chromosome 20;

KW haematopoietic disorder; central nervous system disease; viral infection;

KW peripheral nervous system disease; non-healing wound; infectious disease;

KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;

KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;

KW antiallergic; antiinflammatory; immunosuppressive neuroprotective;

cytostatic; haemostatic; virucide; antibacterial; fungicide;  
immunostimulant; cerebroprotective; gene therapy; gene; ss.

Homo sapiens.

WO200259260-A2.

01-AUG-2002.

16-NOV-2001; 2001WO-US042950.

17-NOV-2000; 2000US-00714936.

(HYSE-) HYSEQ INC.

Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

WPI; 2002-590824/63.  
N-PSTDB; ABP64788.

New isolated polynucleotide, useful in research, diagnostic or  
therapeutic methods, e.g. preventing or treating disorders involving  
aberrant protein expression or biological activity.

Claim 1; SEQ ID NO 107; 394pp; English.

The present invention relates to novel human coding sequences (ABQ99268-  
ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in  
therapeutic, diagnostic and research methods. The polynucleotides may be  
used in the field of molecular biology as hybridisation probes, primers  
for PCR, for chromosome and gene mapping, for the recombinant production  
of protein, or in generation of anti-sense DNA or RNA. The  
polynucleotides are useful in diagnostics as expressed sequence tags  
(ESTs) for identifying expressed genes or for physical mapping of the  
human genome. The proteins may be used as molecular weight markers, or as  
nutritional sources or supplements. The proteins may be used to maintain  
and expand cell population in a totipotent or pluripotential state,  
useful for re-engineering damaged or diseased tissues, transplantation,  
manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
polynucleotides and proteins are useful for preventing, treating or  
ameliorating disorders involving aberrant protein expression or  
biological activity, e.g. haematopoietic disorders, central/peripheral  
nervous system diseases, mechanical and traumatic disorders, non-healing  
wounds, immune deficiencies and disorders, infectious diseases caused by  
viral, bacterial or fungal infection, autoimmune disorders, allergic  
reactions and conditions, coagulation disorders, or cancer. The  
polynucleotide sequences of the invention were assembled from ESTs  
isolated mainly by sequencing by hybridisation, and in some cases,  
sequences obtained from one or more public databases. Note: The sequence  
data for this patent did not form part of the printed specification, but  
was obtained in electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pst\_sequences

Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:
Score: 2,21e-96	1413
Percent Similarity: 1206.50	Matches: 240
Best Local Similarity: 91.60%	Conservative: 0
Query Match: 89.17%	Mismatches: 4
DB: 6	Indels: 18
	Gaps: 1

US-09-939-853A-75 (1-261) x ABQ99374 (1-1413)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20

Db 54 ATGGAGGTCTGCCACGACGAGAAGAAATCTGCACCAAGCCAGCTTGAGTTCCTCTC 113

QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40

Db 114 CAAGGCCAGGACCTGTGACCATTGAAGCAGAGAAGCAAGCGCACCGCTGCCCTG 171

Query Match: 93.79% Indels: 0

DB: 3 Gaps: 0

US-09-939-853A-75 (1-261) x AAC77202 (1-837)

QY 15 SerLeuSerSerValGlnGlnGlyProValThrMetGluAlaGluArgSerLys 34

Db 4 ACCTTGAGTTCTCTGTCCAGGCCAGGACCTGTGACCATGGGAAGCAGAGAACGAG 63

QY 35 AlaThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeuSerLeuArg 54

Db 64 GCCACAGCGCGCCCTGGCAGTTTCCCGCAGGTGGCCGCGGAGCTGTCTCCTGAGA 123

QY 55 LeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrValLeuSerGlu 74

Db 124 CTCGGGGAGCCATTGACCATCTCTCTGAGGTGGAGACTGGTGACGGTCTGTCTGAA 183

QY 75 ValSerGlyArgGluTyraSnIleProSerValHisValGlyLysValSerHisGlyTrp 94

Db 184 GTCTCAGGCAGAGAGATAAATCATCCCCAGCGTCCACGTGGCAAAGTCTCCCATGGGTGG 243

QY 95 LeuTyrgluGlyLeuSerArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnPro 114

Db 244 CTGTATGAGGGCTGAGCAGGAGGAGAAACAGAGGACCTGTCTGTTACCTGGGAACCT 303

QY 115 GlyGlyAlaPheIleuLeuArgGluSerGlnThrArgArgGlySerTyrrSerLeuSerVal 134

Db 304 GGAGGGGCTTCTCATCGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTCTCAGTC 363

QY 135 ArgLeuSerArgProAlaSerTrpAspArgIleArgHISrtyrArgIleHisCysLeuAsp 154

Db 364 CGCCTCAGCGCCCTGCTCCTCGGACCGGATCAGACACTACAGATCCACTGCTTGAC 423

QY 155 AsnGlyTrpLeuTyrlleSerProArgLeuthrPheProSerLeuGlnAlaLeuValasp 174

Db 424 AATGGCTGGCTGTATCATCCCGCGCTCACCTTCCCCTCACCCAGCCCTGGTGGAC 483

QY 175 HisTyrrSerGluLeuAlaAspAspileCysCysLeuLeuLysGluProCysValleuGln 194

Db 484 CATTACTCTGAGCTGGCGGATCACAATCTGCTGCTTACTCAAGGACCCCTGTGCTGCAG 543

QY 195 ArgAlaGlyProLeuProGlyLysAspileProLeuProValThrValGlnArgThrPro 214

Db 544 AGGCTGGCCGCTCCTCGGAGGATATACCCCTACTGTGATGTGCAGAGACACCA 603

QY 215 LeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGlu 234

Db 604 CTCACCTGAAAAGAGCTGCACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGGAG 663

QY 235 SerLeuLeuSerGluGlyLeuArgGluSerPheTyrrIleSerLeuAsnAspGlu 254

Db 664 TCCTTCTCAGTGAGGGTCTCGGAGATCCCTCAGCTTCTATCAGCTCTGAATGACGAG 723

QY 255 AlavalSerLeuAspAla 261

Db 724 GCTGTCTCTTTGATGATGCC 744

RESULT 6

ABQ99374

ID ABQ99374 standard; cDNA; 1413 BP.

XX AC ABQ99374;

XX DT 25-FEB-2003 (first entry)

XX DE Human coding sequence SEQ ID 107.

XX KW Human; expressed sequence tag; EST; chromosome 20;

KW haematopoietic disorder; central nervous system disease; viral infection;

KW peripheral nervous system disease; non-healing wound; infectious disease;

KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;

KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;

KW antiallergic; antiinflammatory; immunosuppressive neuroprotective;

```

QY 41 GlySerPheProAlaGlyClyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 174 GCAGATTTCCCGGAGTGGCCCGGAGCTGTCTGAGACTCGGGAGCCATTGACC 233
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 234 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGAAGTCTCAGGCAGAGAGTAT 293
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluLysLeuSer 100
Db 294 AACATCCCGAGCGTCACGTGCCAAAGTCTCCCATGGTGGCTGTATGAGGGGCTGAGC 353
QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuIle 120
Db 354 AGGGAGAAAGCAGAGAACTGCTGTGTACCTGGGAACCCCTGGAGGGGCTTCTCATC 413
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 414 CGGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCGCCCTGCA 473
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 474 TCTTGGGACCGATCAGACACTACAGGATCCACTGCTTGCATATGGCTGGCTGTATCATC 533
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 534 TCACCGCGGCTCACCTTCCCTCACTCAGGCGCTGGGGAGCCATTAC----- 581
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200
Db 582 -----TCTGAGGCTGGCCCGCTCCC 602
QY 200 cGlyLysAspIleProLeuProValThrValcInArgThrProLeuAsnTrpLysGluLe 220
Db 603 TGGCAAGGATATACCCCTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662
QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGluGlu 240
Db 663 GGACAGCTCCCTCCCTGTTTCTGAAAGCTGCCACAGGGAGGAGTCTTCTTCAGTGAGGG 722
QY 240 yLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAs 260
Db 723 TCTCGGGAGTCCCTCAGCTTCTACATCAGGCTGAATCAGGAGGCTGTCTCTTTGGATGA 782
QY 260 pAla 261
Db 783 TGCC 786

RESULT 7
ID AAL44090
XX AAL44090 standard; cDNA; 737 BP.
XX
AC AAL44090;
XX
DT 03-OCT-2002 (first entry)
XX
DE Mouse MARS short isoform protein coding sequence.
XX
KW Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer.
XX
OS Mus sp.
XX
FH Location/Qualifiers
FT 1..533
FT /*tag= a
FT /product= "Mouse MARS short isoform protein"
XX
PN WO200242452-A2.
XX

```

```

PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-CA001662.
XX
PR 27-NOV-2000; 2000CA-02324663.
XX
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Mcglade JC, Loreto MP;
XX
DR WPI; 2002-566564/60.
DR P-PSDB; AAO15458.
XX
PT New isolated modulator of antigen receptor signaling protein or its
PT fragment, useful for treating malignant disorders such as myeloid
PT malignancies, autoimmune disorders and myeloproliferative disorders.
XX
XX Claim 9; Page 77; 110pp; English.
XX
CC The invention comprises the amino acid and coding sequences of modulator
CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC putative tumour suppressor gene and exhibits structural and sequence
CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
CC protein sequences of the invention are useful for the treatment of
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC disorders, immunosuppression, myeloproliferative disorders and
CC malignancies related to the de-regulation of tyrosine kinases (e.g.
CC breast cancer). The present cDNA sequence encodes a mouse MARS protein.
XX
SQ Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 7,25e-96 Length: 737
Score: 1196.50 Matches: 240
Percent Similarity: 92.02% Conservative: 2
Best Local Similarity: 91.25% Mismatches: 2
Query Match: 88.43% Indels: 19
DB: 6 Gaps: 1
US-09-939-853A-75 (1-261) x AAL44090 (1-737)
QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal 20
Db 1 ATGGGAAGTCTGCCAGCAGCAGAAATACTCTGCCAAGCCCAAGCTTGAGTTCTCTCTGTC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGGCCAGGGACCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGTGGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 121 GGCAGTTTCCCGCAGGTGGCCCGCAGCTGTCTGAGACTCGGGAGGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCGTCTCTGAGATGGAGACTGGTGGACGGTGTCTGTAAGTCTCAGGCAGAGAGTAT 240
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluLysLeuSer 100
Db 241 AACATCCCGAGGTCACGTGCCAAAGTCTCCCATGGTGGCTGTATGAGGGGCTGAGC 300
QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuIle 120
Db 301 AGGGAGAAAGCAGAGAACTGCTGTGTATTACCTGGGAACCCCTGGAGGGGCTTCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGGAGAGCCAGACAGGAGGCTCTTACTCTGTCTGCTCCGCTCAGCCGCCCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGTGACATGGCTGGCTGTACATC 480

```

QY 161 SerProArgLeuThrPheProSerLeuGluAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 481 TCACCGCCCTCACCTCCCTCACTCCAGCCCTGGTGGACCATAC----- 528  
 QY 181 AspAspIleCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200  
 Db 529 -----TCTGAGGCTGGCCGCTCC 549  
 QY 200 oGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220  
 Db 550 TGGCAAGGATATACCCCTACCTGACTGTCCGAGGACACCACTCACTGGAAGAGCT 609  
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGl 240  
 Db 610 GCACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTAGTGAGGG 669  
 QY 240 YLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA 260  
 Db 670 TCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAAATGAGCGAGGCTGTCTCTTTGGATG 729  
 QY 260 spAla 261  
 Db 730 ATGCC 734  
 RESULT 8  
 AAS74750  
 ID AAS74750 standard; cDNA; 2049 BP.  
 AC AAS74750;  
 DT 13-FEB-2002 (first entry)  
 DE DNA encoding novel human diagnostic protein #10554.  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 OS Homo sapiens.  
 QX WO200175067-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 QX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI: 2001-639362/73.  
 DR P-PSDB; ABG10563.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 QX Claim 1; SEQ ID NO 10554; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostic, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;  
 Alignment Scores:  
 Pred. No.: 2,52e-89 Length: 2049  
 Score: 1128.50 Matches: 238  
 Percent Similarity: 89.51% Conservative: 1  
 Best Local Similarity: 89.14% Mismatches: 5  
 Query Match: 89.41% Indels: 23  
 DB: Gaps: 1  
 US-09-939-853A-75 (1-261) x AAS74750 (1-2049)  
 QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 Db 965 ATGGGAAGTCTCCCGAGGAGAGAAATCTTCCAGCCCAAGCTTGAGTTCTCTGTC 1024  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 1025 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCGTGGCCCTG 1084  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 1085 GGCAGTTTCCCGCAGGTGGCCCGGAGCTGCTCGAGACTCGGGAGGACATGACC 1144  
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 1145 ATCTCTCTGAGATGGAGACTGGTGGACGGTGTCTGAAGTCTCAGGCAGAGATAT 1204  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 Db 1205 AACATCCCGCAGGTCACGTGGCCAAAGTCTCCATGGTGGCTGTATGAGGGCTGAGC 1264  
 QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuIle 120  
 Db 1265 AGGAGAGACAGAGAGAACTGCTGTGTACTTGGAAACCTTGGAGGGCCCTTCTCATC 1324  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 1325 CGGAGAGCCAGACCCAGGAGAGGCTCTTACTCTCTGTCTGCTCCGCTCAACGCCCTGCA 1384  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 Db 1385 TTCTGGACCGGATCAGACATCAGGATTCATGGCTTGACATGGCTGGCTGTACATC 1444  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 1445 TCACCGGCTCACCTTCCCTCACTCCAGGCTTGGTGGACCATAC----- 1492  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200  
 Db 1493 -----TCTGAGGCTGGCCGCTCC 1513  
 QY 200 oGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220  
 Db 1514 TGGCAAGGATATACCCCTACCTGACTGTCCAGAGGACACCACTCACTGGAAGAGCT 1573  
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGluGluSerLeuLeuSerGlu 239  
 Db 1574 GGACAGTCCCTCCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGAGGAG 1633  
 QY 240 GlyLeuArgGluSer-LeuSerPheTyrIleSer-LeuAsnAspGluAlaValSer-Leu 258

Db 1634 GGGCTCGGAGATCCCTCAGTCTCTACATCAGCCCTGAATGACGAGGCTGTCTCTTTTG 1693

Qy 259 AspAspAla 261  
|||||

Db 1694 GATGATGCC 1702

RESULT 9

AAAL44087

ID AAL44087 standard; cDNA; 1348 BP.

XX

AC AAL44087;

XX

XX 03-OCT-2002 (first entry)

XX Mouse modulator of antigen receptor signalling protein coding sequence.

DE

DE

XX Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;

KW MARS; tumour suppressor gene; Src-like adaptor protein; SLAP;

KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;

KW immunosuppression; myeloproliferative disorder; breast cancer.

XX

XX Mus sp.

OS

XX

XX Location/Qualifiers

Key 282..1061

FT CDS

FT /\*tag= a

FT /product= "Mouse MARS protein"

XX

XX W0200242452-A2.

PN

XX

XX 30-MAY-2002.

XX

XX 26-NOV-2001; 2001WO-CA001662.

PF

XX

XX 27-NOV-2000; 2000CA-02324663.

PR

XX

XX (HOSP-) HOSPITAL FOR SICK CHILDREN.

PA

XX

XX McGlade JC, Loreto MP;

PI

XX

XX WPI; 2002-566564/60.

DR

XX P-PSDB; AAO15456.

XX

XX New isolated modulator of antigen receptor signaling protein or its

PT fragment, useful for treating malignant disorders such as myeloid

PT malignancies, autoimmune disorders and myeloproliferative disorders.

PT

XX

XX Claim 10; Fig 1A; 110pp; English.

PS

XX

XX The invention comprises the amino acid and coding sequences of modulator

CC of antigen receptor signalling (MARS) proteins. The MARS protein is a

CC putative tumour suppressor gene and exhibits structural and sequence

CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and

CC protein sequences of the invention are useful for the treatment of

CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune

CC disorders, immunosuppression, myeloproliferative disorders and

CC malignancies related to the de-regulation of tyrosine kinases (e.g.

CC breast cancer). The present cDNA sequence encodes a mouse MARS protein

CC

XX

XX Sequence 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.: 9.84e-81 Length: 1348

Score: 1028.00 Matches: 208

Percent Similarity: 85.50% Conservative: 16

Best Local Similarity: 79.39% Mismatches: 34

Query Match: 75.98% Indels: 4

DB: 6 Gaps: 3

US-09-939-853A-75 (1-261) x AAL44087 (1-1348)

Qy 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
|||||

282 ATGGGAAGTTTGTCCAGACAGAGGGAACCC---TCAGCCCCCAGCCCGCCAGCTCCTCTGGT 338  
21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40  
339 CCAGACGAGAACCCCTGTCCATGCAACGAGAAGACACAAGGTCCAGCTGTGGGCCCTG 398  
41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
399 GGCAGTTTCCACAGGTGACAGGCGACACTATCTCTGAGACTCGGGAGCCGCTGACC 458  
61 IleValSerGluAspGlyAspThrPtpThrValLeuSerGluValSerGlyArgGluTyr 80  
459 ATCATCTCTCAGGATGGAGATTGCTGGACAGTCCAGTCGGAAGTCTCAGGCACAGAGTAC 518  
81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
519 CACATGCCAGTGTGTATGTGGCTAAAGTCGCCACGGGTGGTGTATCGAGGGCCCTGAGC 578  
101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
579 CGGAGAGAAACCCAGAGAACTACTTCCTGTTACTTGGGAACCCCGAGGGGCTTTCCTCATC 638  
121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
639 CGGAGAGCCAGACACAGGAGAGGCTGCTATTCCCTGCTCCGTCGAGCTCAGCCGCCCTGCA 698  
141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
699 TCTTGGACCGGATCAGACACTACAGGATACAGCGCTTGACAAATGGCTGGCTGTATCATC 758  
161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
759 TCACCTCGCCTCACCTTCCCTCACTCCAGCCCTTGGTGGAGCATTACTCTGAGCTAGCA 818  
181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
819 GATGGCATCTGCTGCTCCCTCAGGAGCCGCTGTCTCTGCAGAAGCTTGGCCACATCACT 878  
201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
879 GGCAGAAGATACCTCCACCTGTGACTGTGCGCAACATCATCACTAAATTGGAAAAAGCTG 938  
221 AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuLeuSerGlu 239  
939 GACCGCAGCCTCTCTGTTCTGGAAGACCTCGCAGTGGGGAGGCATCTCTGCTCAGTGAG 998  
240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259  
999 GGGCTCCGAGAGTCCCTCAGTTCTCTACATCAGCTCGCTGAGGAC-----CCCTGGAT 1052  
260 AspAla 261  
1053 GATGCT 1058  
RESULT 10  
ABQ98670  
ID ABQ98670 standard; DNA; 763 BP.  
XX AC ABQ98670;  
XX DT  
XX 04-NOV-2002 (first entry)  
XX DE  
XX Human ORF477 coding sequence.  
XX Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
XX Antinflammatory; gene therapy; human; ORPX; atherogenic; platelet;  
XX human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
XX cancer; cardiovascular disease; allergy; autoimmune disease; ds.  
XX wound healing; blood coagulation disorder; inflammatory disorder; ds.  
XX Homo sapiens.  
XX OS  
XX PN US2002082206-A1.

XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX PR 30-MAY-2000; 2000US-0208427P.  
 XX PA (LEAC/) LEACH M D.  
 XX PA (MEHR/) MEHRABAN F.  
 XX PA (CONL/) CONLEY P B.  
 XX PA (TOPP/) TOPPER J N.  
 XX PA (LAWD/) LAW D.  
 XX PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX DR WPI; 2002-626554/67.  
 XX DR P-PSDB; ABP64107.  
 XX PT New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 XX PS Claim 2; SEQ ID NO 953; 78pp; English.  
 XX CC The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?docID=20020082206  
 XX SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,59e-63 Length: 763  
 Score: 826.00 Matches: 158  
 Percent Similarity: 99.37% Conservative: 0  
 Best Local Similarity: 99.37% Mismatches: 1  
 Query Match: 61.05% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-939-853A-75 (1-261) x ABQ98670 (1-763)  
 QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
 DB 286 ATGGGAAGTCTGCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 345  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 DB 346 CAAAGCCAGGACCTGTACCATGTGAACAGAGAGAACAGGACGCGTGGCCCTG 405  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGlyGluProLeuThr 60  
 DB 406 GGCAGTTTCCCGCAGGTGGCCCGCCGAGTGTCTGAGACTCGGGAGCCATTGACC 465  
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 466 ATCGTCTCTGAGATGGAGACTGGTGGACGGTGTCTGAGTCTCAGGCAGAGAGTAT 525  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
 DB 526 AACATCCCGAGGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCTCAGC 585  
 QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuIle 120

DB 586 AGGAGAAAGCAGAGAGAACTGCTGTTTACCTGGGAACCCCTGGAGGGGCTTCTCATC 645  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 646 CGGAGAGCCAGAGCCAGGAGAGGCTCTTACTCTGTCTGTCAGTCCGCTCAGCGCCCTGCA 705  
 QY 141 SerTrpAspArgGileArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyr 159  
 DB 706 TCTTGGGACCGATCAGACACTACAGGATCCACTGCTTGCATATGCTGGCTGTATC 762  
 RESULT 11  
 AAS74748  
 ID AAS74748 standard; cDNA; 603 BP.  
 XX AC AAS74748;  
 XX DT 13-FEB-2002 (first entry)  
 XX DE DNA encoding novel human diagnostic protein #10552.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR P-PSDB; ABG10561.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PS Claim 1; SEQ ID NO 10552; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;  
 Alignment Scores:

Pred. No.: 4.2e-55 Length: 603  
 Score: 731.00 Matches: 151  
 Percent Similarity: 82.45% Conservative: 4  
 Best Local Similarity: 80.32% Mismatches: 21  
 Query Match: 54.03% Indels: 12  
 DB: 5 Gaps: 4

US-09-939-853A-75 (1-261) x AAS74748 (1-603)  
 QY 78 ArgGluTyrAsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGlu 97  
 Db 61 CGCAGGGGAAACCTCCCAAGAGTCGGTG---AAGATCCTCCGGAGCTGGCTGAC--- 114  
 QY 98 GlyLeuSerArgGlyLysAlaA-----GluGluLeuLeuLeuProGlyAsn 113  
 Db 115 ---TTGCACCCCTACAAGCCTACCCCTCAGAGGAGAGCTGAGCCCTTCTGGACAG 171  
 QY 114 ProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTyrSerLeuSer 133  
 Db 172 ACCAACCTGTGCAGTGCTG-----CAAAAGACATGGCTCTTACTCTCTGTCA 216  
 QY 134 ValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIleHisCysLeu 153  
 Db 217 GTCCSCCTCAGCGCCCTCGATCTCTGGACCGGATCAGACACTCAGATCCATGCGCTT 276  
 QY 154 AspAsnGlyTyrLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuVal 173  
 Db 277 GACATGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCCTCAGCGCCCTGGTG 336  
 QY 174 AspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysValLeu 193  
 Db 337 GACCAATTACTCTGAGCTGGCGATGACATCTGCTTACTCAAGGAGCCCTGTGTCTGT 396  
 QY 194 GlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThr 213  
 Db 397 CAGAGGGTGGCCCGCTCCCTGGCAAGATATACCCCTACCTGTGACTGTGCAGAGACA 456  
 QY 214 ProLeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGlu 233  
 Db 457 CCACCTCAATCGAAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAG 516  
 QY 234 GluSerLeuLeuSerGlyGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp 253  
 Db 517 GAGTCTCTCTCAGTGAGGGTCTCGGAGTCCCTCAGCTTCTACATCAGCCTGATGAC 576  
 QY 254 GluAlaValSerLeuAspAla 261  
 Db 577 GAGGCTGTCTCTTGGATGATGCC 600

RESULT 12  
 ABQ99151  
 ID ABQ99151 standard; DNA; 875 BP.  
 XX AC ABQ99151;  
 XX DT 04-NOV-2002 (first entry)  
 XX DE Human ORF958 coding sequence.  
 XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 XX OS Homo sapiens.  
 XX PN US2002082206-A1.  
 XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX ID ABX62975 standard; cDNA; 3756 BP.

PR 30-MAY-2000; 2000US-0208427P.  
 XX (LEAC/) LEACH M D.  
 PA (MEHR/) MEHRABAN F.  
 PA (CONL/) CONLEY P B.  
 PA (TOPP/) TOPPER J N.  
 PA (LAWD/) LAW D.  
 XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 PI WPI; 2002-626554/67.  
 XX P-PSDB; ABP64588.  
 DR New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 XX Claim 2; SEQ ID NO 1915; 78pp; English.  
 PS The present invention relates to novel human ORFX polypeptides and their  
 XX coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
 XX SQ Sequence 875 BP; 205 A; 259 C; 225 G; 185 T; 0 U; 1 Other;

Alignment Scores:  
 Pred. No.: 3.66e-42 Length: 875  
 Score: 586.00 Matches: 112  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 43.31% Indels: 0  
 DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x ABQ99151 (1-875)  
 QY 150 IleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeuThrPheProSerLeu 169  
 Db 4 ATCCACTGGCTTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCCTC 63  
 QY 170 GlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLysGlu 189  
 Db 64 CAGGCCCTGGTGACCATTTACTCTGAGCTGGCGGATGACATCTGCTGCTATCAGGAG 123  
 QY 190 ProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThr 209  
 Db 124 CCCTGTGTCTCTGACAGAGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACTGTGACT 183  
 QY 210 ValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAla 229  
 Db 184 GTGCAGAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTTGAGCT 243  
 QY 230 AlaThrGlyGluGluSerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIle 249  
 Db 244 GCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATC 303  
 QY 250 SerLeuAsnAspGluAlaValSerLeuAspAla 261  
 Db 304 AGCCTGATGACGAGGCTGTCTCTTTGGATGATGCC 339

RESULT 13  
 ABX62975  
 ID ABX62975 standard; cDNA; 3756 BP.



XX AC ABX62975;  
 XX DT 25-FEB-2003 (first entry)  
 XX DE Human activated T cell cDNA #91.  
 XX KW T cell; gene; ss; differential expression; T cell activation;  
 KW antiallergic; cytostatic; immunosuppressive; antimicrobial; gene therapy;  
 KW allergy; cancer; graft versus host disease; infection;  
 KW autoimmune disorder.  
 XX OS Homo sapiens.  
 XX PN US2002137077-A1.  
 XX PD 26-SEP-2002.  
 XX PF 25-OCT-2001; 2001US-00002600.  
 XX PR 25-OCT-2000; 2000US-0243521P.  
 XX PA (HOPK/) HOPKINS C M.  
 XX PA (PETE/) PETERSON D P.  
 XX PA (COCK/) COCKS B G.  
 XX PA (HAWK/) HAWKINS P R.  
 XX PI Hopkins CM, Peterson DP, Cocks BG, Hawkins PR;  
 XX DR WPI; 2003-102381/09.  
 XX PT New combination comprising several cDNAs that are differentially  
 PT expressed in activated T cells, useful for diagnosing, treating, staging  
 PT or monitoring treatment for allergy, cancer, infectious and/or autoimmune  
 PT disorders.  
 XX PS Claim 1; Page: 180pp; English.  
 CC This invention relates to the sequences of several cDNAs that are  
 CC differentially expressed in activated T cells. The sequences of the  
 CC invention may have antiallergic, cytostatic, immunosuppressive and  
 CC antimicrobial activity and may be used in gene therapy. The invention  
 CC also comprises a method for screening samples for differentially  
 CC expressed genes and a method for detecting these cDNAs by hybridisation.  
 CC The methods and compositions of the present invention are useful for  
 CC diagnosing, treating, staging or monitoring treatment for allergy,  
 CC cancer, chronic graft versus host disease, infectious and/or autoimmune  
 CC disorders. The present sequence represents a cDNA of the invention that  
 CC is differentially expressed in activated T cells  
 XX SQ Sequence 3756 BP; 1007 A; 918 C; 960 G; 871 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9.58e-33 Length: 3756  
 Score: 488.00 Matches: 101  
 Percent Similarity: 57.20% Conservative: 46  
 Best Local Similarity: 39.30% Mismatches: 94  
 Query Match: 36.07% Indels: 16  
 Gaps: 3

US-09-939-853A-75 (1-261) x ABX62975 (1-3756)

Qy 5 ProSerArgArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24  
 Db 1098 CCAGGAAAGAAAGAAATGGGAACAGCATGAATCCACCCCTGCGCTGCGGAGG 1157

Qy 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44  
 Db 1158 CCCCTGCCCAACCCGAGGAGGACTGGATAGCGACTTCCTTCGCTGCTAAGTACTACCCG 1217

Qy 45 AlaGlyGlyProAlaGluLeuSerLeuArgGluGlyGluProLeuThrIleValSerGlu 64  
 Db 1218 TCTCTGACATCAGCCCCCGATATTCGCGGAGGGAGAAATGCTGTGATTTCTGAT 1277

Qy 65 AspGlyAspTrrTrrValLeuSerGluValSerGlyArgGluTyrAsnIleProSer 84  
 Db 1278 GAAGGGGCTGGTGGAAAGCTATTCTTAGCACTGGTGGAGAGAGTACATCCCTGGA 1337

Qy 85 ValHisValGlyLysValSerHisGlyTrrLeuTyrGluGlyLeuSerArgGluValAla 104  
 Db 1338 ATATGTGGCCAGAGTTTACATGGCTGGCTGTTTGGAGGCTGGCCAGAGCAAGGCC 1397

Qy 105 GluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuIleArgGluSerGln 124  
 Db 1398 GAGGAGCTGCTGCAGTCCAGACACAAAGGTCGCTCTTCATGATCAGAGAGTGAG 1457

Qy 125 ThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144  
 Db 1458 ACCAAGAAAGGTTTACTCACTGTCGGTGAGACAAAGG-----CAG 1499

Qy 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTrrLeuTyrIleSerProArgLeu 164  
 Db 1500 GTAAAGCATTTACCGCATTTTCGCTGCTGCCCAACACTGGTACTACATTTCCCGAGGCTC 1559

Qy 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspIleCys 184  
 Db 1560 ACCITTCAGTCCGTCGAGGACCTGGTGAACCACTATTCTGAGTGGCTGTGAGTGGCTGTC 1619

Qy 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
 Db 1620 TGTGTGCTCACCACGCTGCTGACACAAAGCAGCGCTGCCCGCAGCAGTGCAGGCTCC 1679

Qy 205 ProLeuProValThrValGlnArgThrProLeuAsnTrrLysGluLeuAspSerSerLeu 224  
 Db 1680 AGCTCACCTGTCACTTCGCTCAGAAAGACTGTGAGTGGAGGAGTGTCCAGA----- 1733

Qy 225 LeuPheSerGluAlaAlaThrGly-----GluGluSerLeuLeu 237  
 Db 1734 ---CTGAGGAGGACCCCGAGGACAGAGAACCCGCTTGGGGTAGAGAGTCCCTTTTC 1790

Qy 238 SerGluGlyLeuArgGluSerLeuSerPheTrrIleSerLeuAsnAspGlu 254  
 Db 1791 AGCTATGGCTTCGAGAGAGCATTGCCTTTACCTGTCCCTGACCCAGTGAG 1841

RESULT 14  
 AAS02049  
 ID AAS02049 standard; cDNA; 2109 BP.  
 AC AAS02049;  
 XX 16-JUL-2001 (first entry)  
 DT DNA encoding molecule for disease detection and treatment, mddt14.  
 DE Human; mddt14; gene therapy; adenosine deaminase deficiency; ADA;  
 KW severe combined immunodeficiency syndrome; cystic fibrosis; thalassemia;  
 KW familial hypercholesterolaemia; haemophilia; factor VIII; factor IX;  
 KW cancer; cell proliferation; parasite; human retrovirus; HIV; hepatitis B;  
 KW hepatitis C; Candida albicans; Plasmodium falciparum;  
 KW Paracoccidioides brasiliensis; Trypanosoma brasiliensis; ss.  
 XX Homo sapiens.  
 OS WO200123538-A2.  
 PN 05-APR-2001.  
 PD 22-SEP-2000; 2000WO-US026085.  
 PF 28-SEP-1999; 99US-0156565P.  
 PR 30-NOV-1999; 99US-0168197P.  
 XX (INCYTE) INCYTE GENOMICS INC.  
 XX Hodgson DM, Lincoln SE, Russo PD, Spiro PA, Banville SC;  
 PI Bratcher SR, Dufour GE, Cohen HZ, Rosen BH, Shah P, Chalup MS;

Qy	5	ProSerArgArgIySerLeuProSerProSerLeuSerSerSerValGlnGln-Gl	24
Db	389	CCAGGGAAGAAAGAAATGGAAACAGCATGAATCCACCCCTGGCCCTGC	448
Qy	24	YProValThrMetGluAlaGluArgSerLyAlaThrAlaValAlaLeuGlySerPhePr	44
Db	449	CCCTCTCCCAACCCGAGGAGCTGATAGCAGCTTCCTTGGCCGTGTAAGTGACTACCC	508
Qy	44	oAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrThrIleValSerGl	64
Db	509	GTCTCTCTGCATCAGCCCCCGCATATTCGCGCAGGGAGAAATCGGTGTGTGATTCGTA	568
Qy	64	uAspGlyAspTTrpThrValLeuSerGluValSerGlyValArgGluIyryxAsnIleProse	84
Db	569	TGAAGGGGCTGGTGAAGACTATTTCCTTAGCACTGGTCCGAGAGATTACATCCCTGG	628
Qy	84	rValHisValGlyLyysValSerHisGlyTrpLeuTyrgIuGlyLeuSerArgGluIyryxAl	104
Db	629	AATATGTGTGGCCAGAGATTACCATGGCTGGCTGTTCCTTAGGGCCCTGGGCAGACACAGGC	688
Qy	104	aGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluIyryxSerGl	124
Db	689	CGAGGAGCTGCTGCACCTGCCAGACACAAAGGTGCGCTCCTTCATGATTCAGAGAGATGA	748
Qy	124	nThrArgArgGlySerTyrrSerLeuSerValArgLeuSerArgProAlaSerTrpAspAr	144
Db	749	GACCAAGAAAGGGTTTTTACTCATCTGTGGTGAGA-----CACAGGCA	790
Qy	144	gIleArgHisTyrrArgIleHisCysLeuAspAsnGlyTrpLeuTyrrIleSerProArgIle	164
Db	791	GGTAAGCATATCCGCATTTTCGTCTGCCCAACCACTGGTACTACATTTCCCCGAGGCT	850
Qy	164	uThrPheProSerLeuGlnAlaLeuValAspHisTyrrSerGluLeuAlaAspAspIleCy	184
Db	851	CACCTTCAGCTGCTGGAGACCTGGTGAACCACTATTCTGAGTGGCTGATGGCTGTG	910

PR 02-OCT-2000; 2000US-0237295P.  
 PR 02-OCT-2000; 2000US-0237316P.  
 PR 03-OCT-2000; 2000US-0237425P.  
 PR 03-OCT-2000; 2000US-0237598P.  
 PR 03-OCT-2000; 2000US-0237604P.  
 PR 03-OCT-2000; 2000US-0237606P.  
 PR 03-OCT-2000; 2000US-0237608P.  
 PR 01-NOV-2000; 2000US-0244867P.  
 PR 01-NOV-2000; 2000US-0245084P.  
 XX (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 DR WPI; 2002-188264/24.  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.  
 XX  
 PS Claim 1; SEQ ID NO 3526; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an anti-  
 CC neoplastic agent. The method involves exposing cells to a chemical agent  
 CC to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening an  
 CC anti-neoplastic agent, and can be used for producing a product which is  
 CC the data collected with respect to the anti-neoplastic agent as a result  
 CC of M1, and the data is sufficient to convey the chemical structure and/or  
 CC properties of the agent. M1 can be used in the treatment of cancer such  
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms's  
 CC tumour  
 XX  
 SQ Sequence 2665 BP; 736 A; 617 C; 689 G; 623 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 7.55e-33 Length: 2665  
 Score: 487.00 Matches: 101  
 Percent Similarity: 57.20% Conservative: 46  
 Best Local Similarity: 39.30% Mismatches: 94  
 Query Match: 35.93% Indels: 16  
 DB: 6 Gaps: 3

US-09-939-853A-75 (1-261) x ABL65189 (1-2665)

QY 5 ProSerArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGly 24  
 Db 24 CCAGGAAAGAAAGAAATGGAAACAGCATGAATCCACCCCTCGCGCTCCGAGAGG 83  
 QY 25 ProValThrMetGluAlaGluArgSerLysLathAlaValAlaLeuGlySerPhePro 44  
 Db 84 CCCCTGCCCAACCCGAGGAGACTGGATAGCGACTTCCTTGGCTGCTAAGTACTACCCG 143  
 QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu 64  
 Db 144 TCTCTGACATCAGCCCGCATATTCGCCGAGGAGAACTGCGTGTGATTCTGAT 203  
 QY 65 AspGlyAspTyrThrValLeuSerGluValSerGlyArgGlyTyrAsnIleProSer 84  
 Db 204 GAAGGGGGCTGGTGAAGACTATTTCTTACACTGCTCGAGAGAGTTACATCCCTGGA 263  
 QY 85 ValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSerArgGluLysAla 104  
 Db 264 ATATGCTGGCCAGAGTTTACCATGGCTGGCTGTTTGGAGGGCTGGGAGAGCAAGGCC 323

QY 105 GluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGlySerGln 124  
 Db 324 GAGGAGCTGCTGAGCTGCCAGACACAAAGGTGGCTCCTTCATCATCAGAGAGTGAG 383  
 QY 125 ThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTyrAspArg 144  
 Db 384 ACCAAGAAGGGTTTACTACTCTGCGTGAGA-----CACAGGAG 425  
 QY 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIleSerProArgLeu 164  
 Db 426 GTAAAGCATTTACGCATTTTCGCTGCGCAACAACTGGTACTACATTTCCCGAGGCTC 485  
 QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCys 184  
 Db 486 ACCTTCCAGTGGCTGAGACCTGGTGAACCACTATTCTGAGTGGCTGATGGCTGTGC 545  
 QY 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
 Db 546 TGTGTGCTCACCACGCCCTGCTGACACAAAGCAGCGCTGCCACAGCAGTGGGCTCC 605  
 QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLeuAspSerSerLeu 224  
 Db 606 AGCTCACCTGTCCCTTGGCTCAGAAAGACTGTGGACTGGAGAGAGTGTCCAGA----- 659  
 QY 225 LeuPheSerGluAlaAlaThrGly-----GluGluSerLeuLeu 237  
 Db 660 ---CTCAGGAGGACCCCGAGGAAACAGAGAACCCGCTGGGGTAGACGAGTCCCTTTTC 716  
 QY 238 SerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu 254  
 Db 717 AGCTATGGCTTCGAGAGAGCAATTGCTTACCTGTCCCTGACCCAGTGAG 767

Search completed: February 20, 2004, 09:56:47  
 Job time : 377 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:39:38 ; Search time 2518 Seconds  
(without alignments)  
3095.325 Million cell updates/sec

Title: US-09-939-853A-75  
Perfect score: 1353  
Sequence: 1 MGSIPSRKSLPSPSLSSV.....RESLSFYISLNDVSLDDA 261

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-O=/cgn2.1/USPTO.spool.p/US09939853/runat.19022004.145339.24496/app.query.fasta.1.455  
-DB=FAST -OEWMT=FAST -SURFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=500 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09939853 @CGN 1 1 4237 @runat.19022004.145339.24496 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESQRY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpi:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1273	94.1	1002	12	BQ052308 AGENCOURT
2	1183	87.4	1069	12	BQ052468 AGENCOURT
3	1148.5	84.9	1020	12	BQ054281 AGENCOURT
4	1028	76.0	2837	11	AK088672 Mus muscu
5	1024	75.7	2974	11	AK030877 Mus muscu
6	933.5	69.0	926	11	AK020837 Mus muscu
7	878.5	64.9	660	13	BY742155 BY742155
8	867	62.1	1201	9	AL541041 AL541041
9	711	52.5	566	12	BQ284179 602408226
10	680.5	50.3	986	12	BQ054265 AGENCOURT
11	611.5	45.2	960	13	BQ944126 AGENCOURT
12	560.5	41.4	878	12	BQ053486 AGENCOURT
13	557	41.2	377	9	AA959151 v251906.r
14	556.5	41.1	660	10	BB635615 BB635615
15	555	41.0	597	9	AL844311 AL844311
16	546	40.4	778	12	BG178487 602328305
17	545	40.3	606	9	AL844309 AL844309
18	543	40.1	614	9	AL844307 AL844307
19	531	39.2	569	14	CB426333 601508 MA
20	517.5	38.2	627	10	BB619854 BB619854
21	492	36.4	2810	11	AK036167 Mus muscu
22	491.5	36.3	831	29	AY416279 Mus muscu
23	491	36.3	1386	11	AK041565 Mus muscu
24	489.5	35.2	831	29	AY416277 Homo sapi
25	478	35.3	1201	9	AL539427 AL539427
26	477.5	35.3	655	14	CF913437 CF913437
27	476.5	35.2	1997	11	AK037901 AK037901
28	473	35.0	1201	9	AL551370 AL551370
29	470.5	34.8	701	13	BX849096 BX849096
30	468	34.6	1106	13	BX436423 BX436423
31	467	34.5	972	13	BQ707614 AGENCOURT
32	464	34.3	1133	9	AL549826 AL549826
33	452.5	33.4	871	13	BQ436143 AGENCOURT
34	447.5	33.1	775	14	CB938697 IDCGJX13
35	442	32.7	1063	13	BQ072745 AGENCOURT
36	441	32.6	723	12	BM950089 UI-M-BHOP
37	439.5	32.5	777	14	CB938723 IPCGJX13
38	439.5	32.5	1003	13	EX415149 EX415149
39	434	32.1	792	14	CA366437 641999 NC
40	433	32.0	767	14	CF287401 AGENCOURT
41	433	32.0	794	12	BG677567 602624118
42	433	32.0	849	12	BI769183 603053793
43	425.5	31.4	603	13	BQ553005 H4019E02
44	424	31.3	570	9	AI471720 tal6a01.x
45	422	31.2	673	10	BB638252 BB638252

# ALIGNMENTS

RESULT 1  
BQ052308  
LOCUS BQ052308  
DEFINITION AGENCOURT\_6868571 NIH\_MGC\_106 Homo sapiens CDNA clone IMAGE:5933542  
5', mRNA sequence.  
ACCESSION BQ052308  
VERSION BQ052308.1 GI:19811648  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1002)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>;  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2118 row: d column: 23  
 High quality sequence stop: 670.

FEATURES  
 source  
 1..1002  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5933542"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 106"  
 /note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site: 2:  
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,246-118 Length: 1002  
 Score: 1273.00 Matches: 256  
 Percent Similarity: 97.36% Conservative: 2  
 Best Local Similarity: 96.60% Mismatches: 3  
 Query Match: 94.09% Indels: 4  
 DB: 12 Gaps: 0

US-09-939-853A-75 (1-261) x BQ052308 (1-1002)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 DB 96 ATGGGAAGTCTCCAGAGAGAAATCTCTGCCAGCCCAAGCTTGAGTTCTCTGTC 155  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAla 40  
 DB 156 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCGCTG 215  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyProLeuThr 60  
 DB 216 GGCAGTTTCCCGCAGGTGCGCCGCGCAGCTGCTGAGACTCGGGAGCCATTGACC 275  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 276 ATCGTCTCTGAGATGGAGACTGGTGACGGTGTCTGAGTCTCAGGCAGAGATAT 335  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 DB 336 AACATCCCGCAGGTCCACGTGCGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 395  
 QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyValAlaPheLeuIle 120  
 DB 396 AGGAGAAAGCAGAGAACTGTGTGTGTATCTGGACCCCTGGAGGGCCCTTCTCATC 455  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 456 CCGGAGAGCAGACCAAGGAGAGCTTACTCTCTGTGTCAGTCCGCTCAGCGCCCTGCA 515  
 QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160

DB 516 TCCTGGACCGGATCAGACACTACAGGATCTACTGCCTTGACATGGCTGGCTGTACATC 575  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 DB 576 TCACCGCGCTCACCTTCCTCTCACTCAGGCGCTGGTGGACCATTAATCTTGAGCTGGCG 635  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 DB 636 GATGACATCTGTGCTACTCAAGAGCGCTGTGCTCTGCAGAGGCTGGCCCGTCCCT 695  
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 DB 696 GGCACAGGATATACCTCTACTGTGACAGGACACCACTCACTGGAAGAGCTG 755  
 QY 221 AspSerSerLeuLeuPheSerGluAlaThr-GlyGluGluSerLeuLeuSerGlu-G 240  
 DB 756 GACAGCTCCTCTCTGTTTCTGAACTGCCACAGGGGAGGAGTCTCTTCAGTGAGGG 815  
 QY 240 LysLeuArg-GluSerLeuSerPheTyrIleSerLeuAsnAspGlu-AlaValSerLeuAs 259  
 DB 816 GTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCCTGATGACGAGGCTGTCTCTTGA 875  
 QY 259 pAspAla 261  
 DB 876 TGAIGCC 882

RESULT 2  
 BQ052468  
 LOCUS AGENCOURT 6868422 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5933772  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BQ052468  
 VERSION BQ052468.1 GI:19811808  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 1069)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2118 row: n column: 13  
 High quality sequence stop: 681.

FEATURES  
 source  
 1..1069  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5933772"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 106"  
 /note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site: 2:  
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

```

Pred. No.: 3,47e-109 Length: 1069
Score: 1193.00 Matches: 243
Percent Similarity: 92.05% Conservative: 0
Best Local Similarity: 92.05% Mismatches: 2
Query Match: 87.44% Indels: 20
DB: 12 Gaps: 1

US-09-939-853A-75 (1-261) x BQ052468 (1-1069)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 89 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC 148
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluAqcSerLysAlaThrAlaValAlaLeu 40
Db 149 CAGGGCAGGAGCTGTGACCATGGAGCAGAGAGAGCAAGCCACACCCGTGGCCCTG 208
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 209 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTCTGCTGAGACTCGGGGAGCATTGACC 268
QY 61 IleValSerGluAspGlyAspTTPThrValLeuSerGluValSerGlyArgGluTyr 80
Db 269 ATCGTCTCTGAGGATGGAGACTGGTGAGCGGTCTCTCTGAAGTCTCAGGCGAGAGTAT 328
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 329 AACATCCCGAGCTCCAGCTGGCCAAAGTCTCCATGGTGTATGAGGGCCCTGAGC 388
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuLeu 120
Db 389 AGGAGAAACACAGAGAACTGCTGTGTTACCTGGGAACCTCGAGGGCCCTTCCTCATC 448
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 449 CGGAGAGCCAGACCCAGGAGAGCTCTTACTCTCTCTGAGTCCGCTCAGCCGCCCTGCA 508
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 509 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTGTTGACAATGGCTGTATCATC 568
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeu-ValAspHisTyrSerGluLeuAl 180
Db 569 TCACCGGGCCCTACCTCCCTCCCTCCTCCAGCCCTGGGTGGACCATCTCTGAGGG--- 625
QY 180 aAspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPr 200
Db 626 -----GCTGGCCCGCTCCC 639
QY 200 oGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220
Db 640 TGGCAAGGATATACCCCTACTCTGACTGTGCGAGGACACCACTCACTGGAAAGAGCT 699
QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGluGluSerLeuSerGluG 240
Db 700 GGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTCTCAGTGAGG 759
QY 240 lyLeuArgGlu-SerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 760 GTCTCCGGGAAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTTGGAT 819
QY 260 AspAla 261
Db 820 GATGCC 825

RESULT 3
BQ054281
LOCUS
DEFINITION AGENCOURT_6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
5', mRNA sequence.
ACCESSION BQ054281
VERSION BQ054281.1 GI:19813621
KEYWORDS EST.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1020) Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/MCI
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2125 row: j column: 11
High quality sequence stop: 556.
FEATURES
Location/Qualifiers
1..1020
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGACAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 1,02e-105 Length: 1020
Score: 1148.50 Matches: 237
Percent Similarity: 93.13% Conservative: 7
Best Local Similarity: 90.46% Mismatches: 11
Query Match: 84.89% Indels: 7
DB: 1 Gaps: 1

US-09-939-853A-75 (1-261) x BQ054281 (1-1020)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 212 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC 271
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 272 CAAGCCAGGAGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCGCTGGCCCTG 331
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 332 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTCTGCTGAGACTCGGGAGCCATTGACC 391
QY 61 IleValSerGluAspGlyAspTTPThrValLeuSerGluValSerGlyArgGluTyr 80
Db 392 ATCGTCTCTGAGGATGGAGACTGGTGGACGTGTCTGTAAGTCTCAGGCAGAGATAT 451
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 452 AACATCCCGAGCTCCAGCTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 511
QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyAlaPheLeuLeu 120
Db 512 AGGAGAAAGCAGAGGAACCTGCTGTGTACCTNGAACCTCGAGGGGGCTTCTCTCATC 571
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140

```

Db 572 CGGAGACCCAGACAGAGGCTCTTACTCTCTGTCAGTCCGCCCTCAGCCGCCCTGCA 631

Qy 141 SerTrpAspArgIleAeGHisTyrArgIleHisCysLeuAspAsnGlyTTPLeuTyrIle 160

Db 632 TCCTGNGACCGGATCAGACACTACAGAGTCCAGCTGCTTGCACATGGCTGGCTGTATCATC 691

Qy 161 SerProArgLeuThrPheProSerLeuGluAla-LeuValAspHisTyrSerGluLeuAl 180

Db 692 TCACCGGCGCTCACCTTCCCTCTACTCCAGGCCCTGTGTGACCATTACTCTGAGCTGGC 751

Qy 180 aAspAspileCysCysLeu-LeuLysGluProCysValLeuGluArgAlaGlyProLeuP 200

Db 752 GGATGACATCTGCTGCTTANCTCAGGACCCCTGTCTGCTGCAAGGCTGCCCCGCTCC 811

Qy 200 roGlyLysAspile-ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGlu 219

Db 812 CTGGCAAGATATACCCCTACCTGTGACTGTGCAAGGACACCATCTACTGGGAGAG 871

Qy 220 LeuAspSerSerLeuLeu---PheSerGluAlaAlaThrGlyGluGluSerLeuLeuSer 238

Db 872 CTGGACAGCTCCCTCTGTTTCTTGAAGCTGCCGCCAGGGGAGGAGTCTCTTCATCATT 931

Qy 239 Gludly--LeuArgGluSerLeuSerPheTyrIleSerLeu-LeuAspGluAla 255

Db 932 GAGGGGTCTCCGGAGTCCCTCAGCTTCTACATCAACCTGTATGACACAGGCT 985

RESULT 4

LOCUS AK088672 2637 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430023D24 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.

ACCESSION AK088672

VERSION AK088672.1 GI:26353729

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PMID 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PMID 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,I., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PMID 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 2637)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gsr.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/

FEATURES

source

1. 2637

Location/Qualifiers

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="NOD"

/db\_xref="FANTOM\_DB:E430023D24"

/db\_xref="MGI:2427569"

/db\_xref="taxon:10090"

/clone="E430023D24"

/cell\_type="thymic cells"

/tissue\_type="thymus"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="2 days neonate"

358..1157

/note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTR)AAL38196, evidence: FASTY, 100%ID, 100%length, match=777)

putative"

/codon\_start=1

/protein\_id="BAC40495.1"

/db\_xref="GI:26353730"

/translation="MGSLSRGKTSRSPSSSGDQBPVSNQPRHKVTAVALGSFPA GEQRLURLGEPLTIISDGDWTVQSEVSGREYHMPVSVVAKVAGWLTGSLUREK AEELLPLNPGCGFLIRESCYRRCYSLSVLRSPASWDRIHYRIQRLDNGLYIS PRLTFFSLHALVEHYSLADGICCPLEPCVLQKLPGLPGKDTPTTPTVPTSSLNWK LDRSLPLLEAPASGEASLLSGLRESLSYSLAEDPLDDA"

ORIGIN

Alignment Scores:

Pred. No.: 8,96e-93 Length: 2637

Score: 1028.00 Matches: 208

Percent Similarity: 85.50% Conservative: 16

Best Local Similarity: 79.39% Mismatches: 34

Query Match: 75.98% Indels: 4

DB: 11 Gaps: 3

```

US-09-939-853A-75 (1-261) x AK088672 (1-2637)
QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
Db 358 ATGGGAAGTTTCTCCAGCAGAGGAAACCC--TCCAGCCCCAGCCAGCTCTCTGCT 414
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40
Db 415 CCAGACAGGAACCCGGTGTCTGACCAACCAAGACACAAAGTCTCACAGTGTGGCCCTG 474
QY 41 GlySerPheProAlaGlyGlyPheProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 475 GGCAGTTTCCAGCAGGTGACAGCCAGACTATCTCTGAGACTCGGGAGCCGCTGACC 534
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 535 ATCATCTCTCAGGATGGAGATTGGTGACAGTCCAGTCCGAAGTCTTCAGCGAGAGTAC 594
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 595 CACATGCCAGTGTGTATGTGGCTAAGTCGCCACGGGTGGTGTACGAGGCGCTGAGC 654
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 655 CGGAGAAAGCCAGGAACCTACTCTGTACCTGGGAACCCCGAGGGGCGCTTCCTCATC 714
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 715 CGGAGAGCCAGACAGGAGAGGCTGCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 774
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 775 TCTTGGACCGGATCAGACACTACAGATACAGCGCTTGACAAATGGCTGGCTGTACATC 834
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 835 TCACCTCGCTCACCCTTCCCTCACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 895 GATGGCATCTGCTCTCCCTCAGGAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpIysGluLeu 220
Db 955 GGCAGAGATACACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
QY 221 AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuLeuSerGlu 239
Db 1015 GACCGAGCCTCTGTTCTTGGAGGACCTGCGAGTGGGGAGGCGATCTCTGCTGCTGCTG 1074
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 1075 GGGCTCCGAGAGTCCCTCAGTTCTTCTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1128
QY 260 AspAla 261
Db 1129 GATGCT 1134

RESULT 5
LOCUS AK030877
DEFINITION Mus musculus adult male thymus cDNA, RIKEN full-length enriched
library, clone:5830437K10 product:MODULATOR OF ANTIGEN RECEPTOR
SIGNALLING MARS, full insert sequence.
ACCESSION AK030877
VERSION AK030877.1 GI:26326848
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
REFERENCE 1

```

Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 99279253  
 10349636  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 20499374  
 11042159  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076861  
 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 685-690 (2001)  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 2974)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/  
 Location/Qualifiers  
 1. 2974  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM\_DB:5830437K10"  
 /db\_xref="MG1:2394046"



```

/db_xref="taxon:10090"
/clone="5830437K10"
/sex="male"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
232..1011
notes="unnamed protein product; MODULATOR OF ANTIGEN
RECEPTOR SIGNALING MARS (SPTR)AAL38196, evidence: FASTY,
100%ID, 100%length, match=777)
putative"
/codon_start=1
/protein_id="BAC27168.1"
/db_xref="GI:26326849"
/translation="MGLSRGRKTSFSPSSSGPDQEPVSMQPERKHTAVALGSFPA
GEARLSRIGEDPTLISEGDWTVQSEVSGREYVYAKVAHGLWLYGLSREK
AEELLPLNPGFALRESQDRTWGYSLVRSRPSMPSYRKHRYIQRLDNGWLYIT
PRITFPLHALVHYSLAGICCPLEPCVLQGLPKQKTPPTVPTVTSLSLWKK
LDRSLLEAPASGEASLLSEGLRESLSYISLAEDPLDDA"
ORIGIN
Alignment Scores:
Pred. No.: 2,86-92 Length: 2974
Score: 1024.00 Matches: 207
Percent Similarity: 85.50% Conservatives: 17
Best Local Similarity: 79.01% Mismatches: 34
Query Match: 75.68% Indels: 4
DB: 11 Gaps: 3
US-09-939-853a-75 (1-261) x AK030877 (1-2974)
QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
DB 232 ATGGGAAGTTGTCCAGCAGAGGGAACCC---TCCAGCCCCAGCCCGACTCTCTGGT 288
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGlnArgSerLySerValAlaAlaLeu 40
DB 289 CCAGACAGGAACCCGTGTCTATCCACCAAGAGACACAGGTCAAGCTGTGGCCCTG 348
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB 349 GCGAGTTTCCAGCAGGTGAACAGCCAGACTATCTTGAGACTCGGGAGCGCTGACC 408
QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGluTyr 80
DB 409 ATCATCTCTGAGGATGAGATTGTGTGACAGTCCAGTCGGAAGTCTCAGCAGAGATC 468
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
DB 469 CACATGCCAGTGTGTATGGCTAAAGTCGCCACCGGTGGCTGTACGAGGCGCTGAGC 528
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
DB 529 CGGAGAAAGCCGAGGAAGTAATCTCTGTATCTGGGAACCCCGAGGGGCTTCTCTCATC 588
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 589 CGGAGAGCCAGACAGAGAGGCTACTATCTCTGTCTCCGACTCAGCCCGCTGCA 648
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspGlnGlyTrpLeuTyrIle 160
DB 649 TCTTGGGACCGGATCAGACACTACAGGATACAGCGTCTTGACAATGCTGGCTGTATCATC 708
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAlaAspHisTyrSerGluLeuAla 180
DB 709 ACACCTGCGCTTACCTTCCCTCTACTCAGCCCTTGTGGAGCATTTACTCTGAGTAGCA 768
QY 181 AspAspIleCysLeuLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200
DB 769 GATGCACTCTGCTGCTCCCTCAGGAGGCGGTGTCTCTCGAGAAGCTTGGCGCACTACCT 828
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
DB 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220

```

```

Db 829 GCGAAGATACACCTCCACCTGTGCTGTGCGCAACATCATCATCAATAATGGAAGCTG 888
QY 221 AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluSerLeuLeuSerGlu 239
Db 889 GACCGACCCCTCTCTGTTTCTGGAAGCAGCTCGAGTGGGAGGCACTCTCTCAGTGAG 948
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 949 GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCTGGCTGGTGGAGGAC-----CCCTTGAT 1002
QY 260 AspAla 261
Db 1003 GATGCT 1008
RESULT 6
AK020837 926 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus adult retina cDNA, RIKEN full-length enriched library,
DEFINITION clone:A930009E21 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING
MARS, full insert sequence.
ACCESSION AK020837
VERSION AK020837.1 GI:12861542
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nishikawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishikawa, K., Katsunari, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Masumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiracka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihata, C., Matsuyama, T., Miyazaki, K., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,

```





QY 260 Asp 260  
 Db 646 GAC 648

AL541041 1201 bp mRNA linear EST 12-MAY-2003  
 AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YK23  
 LOCUS 5-PRIME, mRNA sequence.

ACCESSION AL541041  
 VERSION AL541041.2 GI:30544829  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1201)  
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:12871733.  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 9825.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DE005AF12QP1&cluster=9825.r. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DE005AF12QP1.  
 Location/Qualifiers

FEATURES  
 source  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE005YK23"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with NotI and cloned into  
 the NotI and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,32e-77 Length: 1201  
 Score: 867.00 Matches: 168  
 Percent Similarity: 95.45% Conservative: 0  
 Best Local Similarity: 95.45% Mismatches: 8  
 Query Match: 64.08% Indels: 0  
 DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x AL541041 (1-1201)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 459 ATGGGAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 518  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 519 CAAGCCAGGACCTGTGACCATCGAAGCAGAGAGAGAGCCAGCGGTGGCCCTG 578  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 579 GGCAGATTTCCCGCAGGTGGCCCGCCGAGTGTCTGCTGAGCTCGGGAGCCATTGACC 638  
 QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGlyTyr 80

Db 639 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTGTCTGAAGTCTCAGGCAGAGATAT 698  
 QY 81 AsnIleProSerValHisValGlyValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
 Db 699 AACATCCCGCAGCGTCCAGTCCCAAGTCTCCCATGGTGGCTGTATGAGGGGCTGAGC 758  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeu 120  
 Db 759 AGGGAGAAAGCAGAGAACTGCTGTGTACTGGAAACCCCTGGAGGGGCTCTCTCATC 818  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 819 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTGTCAGCCGCTCAGCGGCTGCA 878  
 QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyr 160  
 Db 879 TCTTGGACCGATCAGACACTTAAAGATCCACTGCTTGCATGGCTGGCTGTACATC 938  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyr 176  
 Db 939 TCACGGCGCTTACCTTCCCTCCCTCAGCGGCTGGTGGACMATTC 986

RESULT 9  
 BG284179 566 bp mRNA linear EST 21-FEB-2001  
 LOCUS 602408226F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4520382 5',  
 DEFINITION mRNA sequence.

ACCESSION BG284179  
 VERSION BG284179.1 GI:13034866  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 566)  
 NIH-MGC http://mgi.mgi.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: c9apbs-remail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10418 row: c column: 07  
 High quality sequence start: 2  
 High quality sequence stop: 566.  
 Location/Qualifiers

FEATURES  
 source  
 1..566  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4520382"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 91"  
 /note="Organ: Prostate; Vector: pCMV-SPORT6; Site: 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.04e-61 Length: 566  
 Score: 711.00 Matches: 136  
 Percent Similarity: 99.27% Conservative: 0  
 Best Local Similarity: 99.27% Mismatches: 1  
 Query Match: 52.55% Indels: 0  
 DB: 12 Gaps: 0

US-09-939-853A-75 (1-261) x BG284179 (1-566)

QY 125 ThrArgGlySerTyrrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144  
 Db 4 ACAGCGTCCGGGCTCTTACTCTGTCTAGTCCGCTCAGCGCTCATCTGGACCGG 63  
 QY 145 IleArgHisTyrrArgIleHisCysLeuAspAsnGlyTrpLeuTyrrIleSerProArgLeu 164  
 Db 64 ATCAGACATACAGATCCACTGCTTGTGAAATGGCTGGCTGTACATCTCACCAGCCCTC 123  
 QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrrSerGluLeuAlaAspAlleCys 184  
 Db 124 ACCTTCCCTCACTCCAGGCTGTGGACCACTTACTCTGAGCTGGCGATGACATCTGC 183  
 QY 185 CysLeuLeuGlyProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
 Db 184 TGCCTACTCAAGAGGCTGTCTCTCCAGAGGCTGGCCGCTCCCTGGCAAGGATATA 243  
 QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224  
 Db 244 CCCCTACTGTGACTGTGCAGAGGACCACTCAACTGGAAGAGCTGGACACTCCCTC 303  
 QY 225 LeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGlyLeuArgGluSer 244  
 Db 304 CTGTTTCTGAAGCTGCCACAGGGAGGAGTCTTCTCAGTGGAGGCTCTCCGGAGTCC 363  
 QY 245 LeuSerPheTyrrIleSerLeuAsnAspGluAlaValSerLeuAspAla 261  
 Db 364 CTCAGCTTCTACATCAGCTGATGACGAGGCTGTCTCTTTGGATGATGCC 414

RESULT 10  
 BQ054265  
 LOCUS  
 DEFINITION AGNCOURT 6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
 5', mRNA sequence.

ACCESSION BQ054265  
 VERSION BQ054265.1 GI:19813605  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 986)  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2125 row: i column: 12  
 High quality sequence stop: 515.

FEATURES  
 source  
 1..986  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5936339"  
 /tissue\_types="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_106"  
 /note="Organ: blood; Vector: pOT7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,35e-58 Length: 986  
 Score: 680.50 Matches: 163  
 Percent Similarity: 73.13% Conservative: 3  
 Best Local Similarity: 71.81% Mismatches: 17  
 Query Match: 50.30% Indels: 44  
 DB: 12 Gaps: 3

US-09-939-853A-75 (1-261) x BQ054265 (1-986)  
 QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 Db 279 ATGGGAAGTCTGCCAGAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCCTCTGTC 338  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 339 CAAGCCAGGACCTGTGACCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGluProLeuThr 60  
 Db 399 GGCAGATTTCCTCCGAGAGTGGCCGCGAGCTGTCTGAGACTCGGGAGCCATTGACC 458  
 QY 61 IleValSerGluAspGlyVasPTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 459 ATCGTCTCTGAGGATGGAGACTGTGGACGGTGTCTGTGAGTCTCAGGAGAGATAT 518  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 Db 519 AACATCCCGCAGCGTCCAGTGGCCAAAGTCTCCCATGGTGGTGTATAGGGGCTGAGC 578  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAla-PheLeuI1 120  
 Db 579 AGGAG 638  
 QY 120 eArgGlu-SerGlnThrArgArgGlySerTyrrSerLeuSerVal-ArgLeuSerArg-Pr 139  
 Db 639 CCGGAG 698  
 QY 139 OAlaSerTrp-AspArg-IleArgHisTyrrArg-IleHisCysLeuAspAsnGlyTrpLe 158  
 Db 699 TGCATCTGGGACCGGAGATCAGACCTCAGAGGATTCCTCCCTGCGCTTGAACCATTTGCT 758  
 QY 158 u-----TyrrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValas 174  
 Db 759 TGGCTTGTAAATTTTAAACCGGGGCTTACCCTTTCCC----- 798  
 QY 174 PHisTyrrSerGluLeuAlaAspAspIleCysLeuLeuLysGluProCysValLeuG1 194  
 Db 799 -----CTTAA 803  
 QY 194 nArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPr 214  
 Db 804 ATTTTCAGGGGCGCTTGGGGGGAACCCCATTT-----ACTCC 839  
 QY 214 cLeuAsn 216  
 Db 840 TTTAAAC 846

RESULT 11  
 BQ944126  
 LOCUS  
 DEFINITION AGNCOURT 10545003 NIH\_MGC\_107 Homo sapiens cDNA clone  
 IMAGE:6728350 5', mRNA sequence.  
 ACCESSION BQ944126  
 VERSION BQ944126.1 GI:24132945  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 960)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM3049 row: m column: 21  
 High quality sequence stop: 628.  
 Location/Qualifiers

FEATURES  
 source

1..960  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6728350"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_107"  
 /notes="Organ: Breast; Vector: pOTB7; Site: 1: EcoRI;  
 Site: 2: XhoI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3.27e-51 Length: 960  
 Score: 611.50 Matches: 124  
 Percent Similarity: 86.71% Conservative: 0  
 Best Local Similarity: 86.71% Mismatches: 1  
 Query Match: 45.20% Indels: 18  
 DB: 13 Gaps: 1

US-09-939-853A-75 (1-261) x BU944126 (1-960)

Qy 120 lIeArgGluSerGlnThrArgArgGlySerTySerLeuSerValArgLeuSerArgPro 139  
 Db 3 ATCCGGAGAGCCACAGCAGGAGGCTTACTCTGTCTGTCAGTCCGCCCTCAGCCGCCCT 62

Qy 140 AlaSerTrpAspArgIleArgHisTyArgIleHisCysLeuAspAsnGlyTrpLeuTy 159  
 Db 63 GCATCTCTGGACCGGATCAGACACTACAGATCCACTGCTTACATGGCTGGCTGTAC 122

Qy 160 lIeSerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTySerGluLeu 179  
 Db 123 ATCTCAGCGGCCCTCAGCTTCCCTCACTCCAGGCCCTGCTGGACCATTA----- 173

Qy 180 AlaAspAspIleCysCysLeuLeuGluProCysValLeuGln-ArgAlaGlyProle 199  
 Db 174 -----TCTGAGGCTGGCCCGCT 191

Qy 199 uProGlyLeuAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpIysG 219  
 Db 192 CCCTGGCAAGATATACCCCTACCTGAGCTGTGCAGAGGACACCACTCACTGGAAGA 251

Qy 219 uLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerG 239  
 Db 252 GCTGGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGCTCTCTCAGTGA 311

Qy 239 uGlyLeuArgGluSerLeuSerPheTyrlIeSerLeuAsnAspGluAlaValSerLeu 259  
 Db 312 GGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCTGANTGACGAGGCTGTCTTTTGA 371

Qy 259 pAspAla 261  
 Db 372 TGATGCC 378

RESULT 12  
 BQ053486  
 LOCUS  
 DEFINITION

BQ053486 878 bp mRNA linear EST 29-MAR-2002  
 AGENCOURT 6822017 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5935253  
 5', mRNA sequence.  
 BQ053486  
 VERSION BQ053486.1 GI:19812826  
 EST.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

1 (bases 1 to 878)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2122 row: l column: 06  
 High quality sequence stop: 394.  
 Location/Qualifiers

1..878  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5935253"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_106"  
 /note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site: 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN

Alignment Scores:  
 Pred. No.: 4.25e-46 Length: 878  
 Score: 560.50 Matches: 134  
 Percent Similarity: 63.60% Conservative: 11  
 Best Local Similarity: 58.77% Mismatches: 21  
 Query Match: 41.43% Indels: 64  
 DB: 12 Gaps: 5

US-09-939-853A-75 (1-261) x BQ053486 (1-878)

Qy 1 MetGlySerLeuProSerArgArgIleSerLeuProSerProSerLeuSerSerVal 20  
 Db 273 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTCTCTCTGTC 332

Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40  
 Db 333 CAAGCCCAAGGACCTGTGACCATGGAAGACAGAGAAAGCCACAGCCGTCGCCCTG 392

Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 393 CGCAGTTTCCCGCAGGCTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 452

Qy 61 lleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80







/clone\_lib="pool YT lib v SPD"

ORIGIN

Alignment Scores: Pred. No.: 7.98e-46 Length: 597  
Score: 555.00 Matches: 107  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 41.02% Indels: 0  
DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x AL844311 (1-597)

QY 155 AsnGlyTrpLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAsp 174  
DB 13 AATGGCTGGCTGTACATCTCACCACGCTCACCTTCCCTCACTCAGGCCCTGGTGGAC 72  
QY 175 HisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysValLeuGln 194  
DB 73 CATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTCAAGGAGCCCTGTCTCTGCAG 132  
QY 195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPro 214  
DB 133 AGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGAGAGGACACCA 192  
QY 215 LeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGlu 234  
DB 193 CTCACCTGGAAGAGCTGGACAGCTCCCTCTCTGTTCTGAGCTGCCACAGGGAGGAG 252  
QY 235 SerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu 254  
DB 253 TCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAG 312  
QY 255 AlaValSerLeuAspAspAla 261  
DB 313 GCTGCTCTTTGGATGATGCC 333

Search completed: February 20, 2004, 11:30:54  
Job time : 2531 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:40:13 ; Search time 89 Seconds  
(without alignments)  
1627.440 Million cell updates/sec

Title: US-09-939-853A-75  
Perfect score: 1353  
Sequence: 1 MSLPERRKSLPSPSLSSV.....RESLSPYISLNDVSLDDA 261

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US0939853/runat\_19022004\_145340\_24508/app\_query.fasta\_1.455  
-DB=Issued Patents NA -QFMR=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -DOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0939853 @CIGN 1.1.85 @runat\_19022004\_145340\_24508 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /cgn2\_6/prodata/2/ina/5A COMB.seq:\*

2: /cgn2\_6/prodata/2/ina/5B COMB.seq:\*

3: /cgn2\_6/prodata/2/ina/6A COMB.seq:\*

4: /cgn2\_6/prodata/2/ina/6B COMB.seq:\*

5: /cgn2\_6/prodata/2/ina/6C COMB.seq:\*

6: /cgn2\_6/prodata/2/ina/6D COMB.seq:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	370.5	27.4	2298	4	US-09-023-655-1158
2	360.5	26.6	2015	4	Sequence 1158, Ap
3	340	25.1	2129	4	Sequence 1105, Ap
4	323	23.9	2435	4	Sequence 1452, Ap
5	320	23.7	2647	4	Sequence 1313, Ap
6	320	23.7	2647	5	Sequence 77, Appl
7	315.5	23.3	4517	4	PCT-US93-06251-77
8	315.5	23.3	4517	5	Sequence 7, Appl
9	313.5	23.2	1491	2	US-09-470-881-7
10	313.5	23.2	1491	3	PCT-US93-06251-83
11	312.5	23.1	2354	4	Sequence 83, Appl
12	289	21.4	1759	4	Sequence 1, Appl
					Sequence 1080, Ap
					Sequence 2, Appl

13	287	21.2	1602	1	US-07-820-011A-1
14	287	21.2	1602	5	PCT-US93-00445-1
15	276.5	20.4	1611	1	Sequence 1, Appl
16	276.5	20.4	1611	4	US-07-820-011A-3
17	276.5	20.4	1611	4	US-09-860-473-3
18	276.5	20.4	1611	5	Sequence 3, Appl
19	262	19.4	1626	4	PCT-US93-00445-3
20	262	19.4	1626	4	US-09-860-473-10
21	243.5	18.0	675	1	Sequence 10, Appl
22	240.5	17.8	2827	4	US-08-707-792A-3
23	240.5	17.8	2827	4	US-08-492-723-1
24	240.5	17.8	2770	4	Sequence 3, Appl
25	240.5	17.8	2770	4	US-08-436-509A-5
26	240.5	17.8	2770	4	Sequence 1, Appl
27	240.5	17.8	2770	4	US-08-232-545-5
28	240.5	17.8	2770	4	Sequence 5, Appl
29	240.5	17.8	2770	4	US-09-006-675-5
30	238	17.6	282	2	Sequence 5, Appl
31	238	17.6	282	3	US-09-238-603A-5
32	230	17.0	1467	4	Sequence 2, Appl
33	230	17.0	1548	4	US-03-579-182-2
34	202	14.9	1661	2	US-09-099-053-1
35	202	14.9	1661	2	US-08-815-176-2
36	193.5	14.3	2187	4	US-09-197-344-2
37	193.5	14.3	2187	4	US-09-023-655-1267
38	190.5	14.1	3623	1	Sequence 1267, Ap
39	182	13.5	1804	1	Sequence 4, Appl
40	182	13.5	1804	1	US-08-306-691B-35
41	182	13.5	1804	4	Sequence 35, Appl
42	172	12.7	874	4	US-08-306-691B-40
43	172	12.7	874	4	Sequence 40, Appl
44	172	12.7	1072	1	Sequence 14, Appl
45	172	12.7	1072	1	PCT-US93-06251-82
					Sequence 82, Appl
					US-09-023-655-931
					Sequence 931, App
					US-07-906-349A-2
					Sequence 2, Appl
					US-08-167-035-5
					Sequence 5, Appl
					US-08-167-035-49
					Sequence 49, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-023-655-1158  
; Sequence 1158, Application US/09023655  
; Patent No. 6607979

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023,655

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: FA-0001 US

; TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1158:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2298 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: g187268  
 US-09-023-655-1158

Alignment Scores:

Pred. No.: 9.15e-30 Length: 2298  
 Score: 370.50 Matches: 80  
 Percent Similarity: 57.71% Conservative: 36  
 Best Local Similarity: 39.80% Mismatches: 76  
 Query Match: 27.38% Indels: 9  
 DB: 4 Gaps: 3

US-09-939-853A-75 (1-261) x US-09-023-655-1158 (1-2298)

QY 6 SerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyPro 25  
 Db 409 TCCATAAACAGCAAGCCAGTTCAGAA---TCTCAGCTTTTACCTGGACAGAGGTTT 465  
 QY 26 ValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerProAla 45  
 Db 466 CAACCTAAGATCCAGAGCAAGAGAGACATCTGTGTAGCCCTGTACCCCTATGATGCC 525  
 QY 46 GlyGlyProAlaGluLeuSerLeuGlyGluProLeuThrLeuValSerGluAasp 65  
 Db 526 ATCCACCGGACGACTTGTCTTTCAGAAAGGAGAGAGATGAAGTCTGTGGAGGACAT 585  
 QY 66 GlyAaspTrpThrValLeuSerGluValSerGlyArgGluTyrAsnIleProSerVal 85  
 Db 586 CGAATGGTGAAGCAAGTCTCTTTTACAAAAAGAGAGCTTCATCCCGACCAAC 645  
 QY 86 HisValGlyLysVal-----SerHisGlyTrpLeuTyrGluGlyLeuSerArg 101  
 Db 646 TATGTGGCCAACTCAACACCTTAGAACACAGAGAGTGGTTTTTCAAGGATATAACACAGG 705  
 QY 102 GluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuIleArg 121  
 Db 706 AAGAGCGCAAGAGCGAGCTTTTGGCCAGGAATAGCGCTGAGGCTTTTCCCTATTAGA 765  
 QY 122 GluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSer 141  
 Db 766 GAAAGTGAACATTAAAGGAAGCTTCTCTGTCTGTCTCAGAGACTTTGACCCCTGTGCAT 825  
 QY 142 TrpAaspGlieArgHisTyrArgIleHisCysLeuAaspGlyTrpLeuTyrIleSer 161  
 Db 826 GGTGATGTTATTAGACACTACAAATATAGAAAGTCTGGATATGGGGCTATTACATCTCT 885  
 QY 162 ProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAasp 181  
 Db 886 CCACGATCACTTTTCCCTGTATCAGCGACATGATTAAACATTACCAAAAGAGGAGAT 945  
 QY 182 AspileCysLeuLeuGlyGluProCysValLeuGlnArgAlaGlyProLeuProGly 201  
 Db 946 GGCCTGTGCAGAAATTCGAGAAGCTTGTATT-----AGTCCCAAGCCACAG 993  
 QY 202 Lys 202  
 Db 994 AAG 996

RESULT 2

US-09-023-655-1105  
 Sequence 1105, Application US/09023655  
 Patent No. 6607879  
 GENERAL INFORMATION:  
 APPLICANT: Cocks, Benjamin G.

APPLICANT: Susan G. Stuart  
 APPLICANT: Jeffrey J. Seilhamer  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
 TITLE OF INVENTION: EXPRESSION  
 NUMBER OF SEQUENCES: 1508  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/023.655  
 FILING DATE: HEREWITH  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0001 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1105:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2015 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: g183911  
 US-09-023-655-1105

Alignment Scores:

Pred. No.: 8.84e-29 Length: 2015  
 Score: 360.50 Matches: 77  
 Percent Similarity: 58.38% Conservative: 31  
 Best Local Similarity: 41.62% Mismatches: 70  
 Query Match: 26.64% Indels: 7  
 DB: 4 Gaps: 2

US-09-939-853A-75 (1-261) x US-09-023-655-1105 (1-2015)

QY 12 ProSerProSerLeuSerSerSerValGlnGlyProValThrMetGluAlaGlu 31  
 Db 286 CCGGGCCCTAATAGCACACAGC-----AACACACAGCAATCAGGAGGCGAGGC 336  
 QY 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeu 51  
 Db 337 TCTGAGGACATCATCGTGTGCCCTGTATGATCAGGAGCCATTCACCAGAGACCTC 396  
 QY 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAaspGlyAspTrpThrVal 71  
 Db 397 AGCTTCCAGAGGGGACACAGATGTTGTCTTAGAGGAATCCGGGAGTGTGGAGGCT 456  
 QY 72 LeuSerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal--- 90  
 Db 457 CGATTCCTGGCCACCCGGAAGAGAGGCTACATCCACAGCAACTATGTCCCGCGTTCAC 516  
 QY 91 -----SerHisGlyTrpLeuTyrGluGlyLeuSerArgGluLysAlaGluLeu 107  
 Db 517 TCTCTGGACAGACAGGAGTGTGTTTTTCAAGGCGCATCAGCCGGAAGAGCGCAGAGCGCAA 576



NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0001 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1313:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2435 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: G338227  
 US-09-023-655-1313

Alignment Scores:  
 Pred. No.: 1,24e-24 Length: 2435  
 Score: 323.00 Matches: 81  
 Percent Similarity: 51.74% Conservative: 23  
 Best Local Similarity: 40.30% Mismatches: 76  
 Query Match: 23.87% Indels: 22  
 DB: 4 Gaps: 3

US-09-939-853A-75 (1-261) x US-09-023-655-1313 (1-2435)

QY 12 ProSerProSerLeuSerSerValGlnGly----- 22  
 DB 507 CCATCCCACTTACMACAACTTCACGCGCGGGGCCAAGGACTCACCGTCTTTGGAG 566  
 QY 23 -----GlnGlyProValThrMetGluAlaGluArgSerLysAla 35  
 DB 567 GTGTGAACCTCTCGTCTCATACGCGGACCTTCGCTACGAGAGGAGAACAGAGTG-ACA 625  
 QY 36 ThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeuSerLeuArgLeu 55  
 DB 626 CTCCTTTGGGCGCTTATAGTACTATGAACAGCTCGGAGAGGATGTTGGTGGAAAGCCGCTCCAT 685  
 QY 56 GlyGluProLeuThrIleVal---SerGluAspGlyAspTrrPrrThrValLeuSerGlu 74  
 DB 686 GGAGAAAAATTTCAAATATTGAACAGCTCGGAGAGGATGTTGGTGGAAAGCCGCTCCAT 745  
 QY 75 ValSerGlyArgGluTyrrAsnIleProSerValHisValGlyLysVal----- 90  
 DB 746 ACAACTGGAGAGACAGGTTACATTTCCAGCAATATATGCTCCAGTTGACTCTATCCAG 805  
 QY 91 SerHisGlyTrrPrrLeuTyrrGluLeuSerArgGluLysAlaGluGluLeuLeu 110  
 DB 806 GCAGAGAGTGGTACTTTGGAAACTTGGCCGAAAGATGCTGAGCGACACTTATGCTCC 865  
 QY 111 ProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTyr 130  
 DB 866 TTGTGAAACCCCAAGAGGTACCTTTCTTATCCGCGAGAGTGAACCAACCAAGGTCCTAT 925  
 QY 131 SerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIle 150  
 DB 926 TCACCTTTCTACCGTCATGCGATGATATGAAGAGACCATGTCACAACTTATAAAAT 985  
 QY 151 HisCysLeuAspAsnGlyTrrPrrLeuTyrrIleSerProArgLeuThrPheProSerLeuGln 170  
 DB 986 CGCAAACTTGACATGTTGGATGATACATTTACCCCGGCGCCAGTTTGAACACTTCAG 1045  
 QY 171 AlaLeuValAspHisTyrSerGluLeuAlaAspIleCysCysLeuLeuLysGluPro 190  
 DB 1046 CAGCTTTGTACAACTTACTTCAGAGAGAGCTGCGAGTCTCTGCTGCCGCTAGTAGTCCC 1105  
 QY 191 Cys 191  
 DB 1106 TGT 1108

RESULT 5

US-09-220-132-77

; Sequence 77, Application US/09220132  
 ; Patent No. 6506607  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shvajan, Andrew W.  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
 ; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER  
 ; FILE REFERENCE: 07334-074001  
 ; CURRENT APPLICATION NUMBER: US/09/220,132  
 ; CURRENT FILING DATE: 1998-12-23  
 ; PRIOR APPLICATION NUMBER: US 60/079,303  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: US 60/069,821  
 ; PRIOR FILING DATE: 1997-12-24  
 ; NUMBER OF SEQ ID NOS: 191  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 77  
 ; LENGTH: 2647  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-220-132-77

Alignment Scores:  
 Pred. No.: 2,97e-24 Length: 2647  
 Score: 320.00 Matches: 80  
 Percent Similarity: 51.74% Conservative: 24  
 Best Local Similarity: 39.80% Mismatches: 76  
 Query Match: 23.65% Indels: 22  
 DB: 4 Gaps: 3

US-09-939-853A-75 (1-261) x US-09-220-132-77 (1-2647)

QY 12 ProSerProSerLeuSerSerValGlnGly----- 22  
 DB 716 CCATCCCACTTACMACAACTTCACGCGCGGGGCCAAGGACTCACCGTCTTTGGAG 775  
 QY 23 -----GlnGlyProValThrMetGluAlaGluArgSerLysAla 35  
 DB 776 GTGTGAACCTCTCGTCTCATACGCGGACCTTCGCTACGAGAGGAGAACAGAGTG-ACA 834  
 QY 36 ThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeuSerLeuArgLeu 55  
 DB 835 CTCCTTTGGGCGCTTATAGTACTATGAACAGCTCGGAGAGGATGTTGGTGGAAAGCCGCTCCAT 894  
 QY 56 GlyGluProLeuThrIleVal---SerGluAspGlyAspTrrPrrThrValLeuSerGlu 74  
 DB 895 GGAGAAAAATTTCAAATATTGAACAGCTCGGAGAGGATGTTGGTGGAAAGCCGCTCCAT 954  
 QY 75 ValSerGlyArgGluTyrrAsnIleProSerValHisValGlyLysVal----- 90  
 DB 955 ACAACTGGAGAGACAGGTTACATTTCCAGCAATATATGCTCCAGTTGACTCTATCCAG 1014  
 QY 91 SerHisGlyTrrPrrLeuTyrrGluLeuSerArgGluLysAlaGluLeuLeuLeu 110  
 DB 1015 GCAGAGAGTGGTACTTTGGAAACTTGGCCGAAAGATGCTGAGCGACAGCTATTTGTC 1074  
 QY 111 ProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTyr 130  
 DB 1075 TTGTGAAACCCCAAGAGGTACCTTTCTTATCCGCGAGAGTGAACCAACCAAGTCCAT 1134  
 QY 131 SerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIle 150  
 DB 1135 TCACCTTTCTACCGTCATGCGATGATATGAAGAGACCATGTCACAACTTATAAAAT 1194  
 QY 151 HisCysLeuAspAsnGlyTrrPrrLeuTyrrIleSerProArgLeuThrPheProSerLeuGln 170  
 DB 1195 CGCAAACTTGACATGTTGGATGATACATTTACCCCGGCGCCAGTTTGAACACTTCAG 1254  
 QY 171 AlaLeuValAspHisTyrSerGluLeuAlaAspIleCysCysLeuLeuLysGluPro 190  
 DB 1255 CAGCTTTGTACAACTTACTTCAGAGAGAGCTGCGAGTCTCTGCTGCCGCTAGTAGTCCC 1314  
 QY 191 Cys 191

Db 1315 TGT 1317

RESULT 6

PCT-US93-06251-77

Sequence 77, Application PC/TUS9306251

GENERAL INFORMATION:

APPLICANT: Wikstrom, Eric and Rife, Jason P.

TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing Stereospecific Alkylphosphonates and Arylphosphonates

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: NY

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/06251

FILING DATE: 19930630

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8586

TELEPHONE: 516-742-4343

TELEFAX: 516-742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 77:

SEQUENCE CHARACTERISTICS:

LENGTH: 2647 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US93-06251-77

Alignment Scores:

Pred. No.: 2,97e-24 Length: 2647

Score: 320.00 Matches: 80

Percent Similarity: 51.74% Conservative: 24

Best Local Similarity: 39.80% Mismatches: 76

Query Match: 23.65% Indels: 22

DB: 5 Gaps: 3

US-09-939-853A-75 (1-261) x PCT-US93-06251-77 (1-2647)

Qy 12 ProSerProSerLeuSerSerValGlnGly----- 22

Db 716 CCATCCCACTACAACTTCCAGCGCGGGCCAGGAGCTCACCCTTTGGAG 775

Qy 23 -----GlnGlyProValThrMetGluAlaGluArgSerLysAla 35

Db 776 GTGTGAACCTTCCTCATCGGAGACCTTCGTACGAGAGGAGGAGACAGAGTG-ACA 834

Qy 36 ThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeu 55

Db 835 CTCTTTGTGCCCCCTTTATGACATATGAGCAGCGGACAGAGATGACCTGATTTTCAAAA 894

Qy 56 GlyGluProLeuThrIleVal-----SerGluAspGlyAspTyrThrValLeuSerGlu 74

Db 895 GGAGAAAAATTTCAAATATTGAACAGCTCGGAGGAGATTGCTGGAGCCCGCTCTTG 954

Qy 75 ValSerGlyArgGluTyrAsnIleProSerValHisValGlyVal----- 90

Db 955 ACAACTGGAGACAGAGTTACATTCACGAATTATGTGGCTCCAGTTGACTCTATCCAG 1014

Qy 91 SerHisGlyTrpLeuTyrGluGlyLeuSerArgGluLysAlaGluLeuLeuLeu 110

Db 1015 GCAGAAGAGTGGTACTTTGGAAAACTTGGCCGAAAGATGCTGAGCGACAGCTATTGTCC 1074

Qy 111 ProGlyAsnProGlyGlyAlaPheLeuLeuLeuSerGlnThrArgArgGlySerTyr 130

Db 1075 TTTGGAACCCCAAGAGTACTTCTTCATCCGAGAGTGAACCCCAAGGTGCCTAT 1134

Qy 131 SerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIle 150

Db 1135 TCACCTTCTATCCGTGATTGGGATGATGAAGAGGACCATGTCAACACATTATAAATT 1194

Qy 151 HisCysLeuAspAsnGlyTyrLeuTyrIleSerProArgLeuThrPheProSerLeuGln 170

Db 1195 CGCAACTTGACAATGTGTGACTACTACATCACCCGGCGCCAGTTTGAACACACTTCAG 1254

Qy 171 AlaLeuValAspHisTyrSerGluLeuAlaAspIleCysCysLeuLysGluPro 190

Db 1255 CAGCTTGTACAACTTACTCAGAGAGAGCTGCAGGTCTCTGCTGCGCCTAGTAGTTCCC 1314

Qy 191 Cys 191

Db 1315 TGT 1317

RESULT 7

US-09-470-881-7

Sequence 7, Application US/09470881

Patent No. 6685938

GENERAL INFORMATION:

APPLICANT: CHERESH, David A.

APPLICANT: ELICEIRI, Brian

TITLE OF INVENTION: METHOGENESIS AND COMPOSITIONS USEFUL FOR MODULATION OF

TITLE OF INVENTION: ANGIOGENESIS AND VASCULAR PERMEABILITY USING SRC OR

TITLE OF INVENTION: YES TYROSINE KINASES

FILE REFERENCE: TSRI 651.2

CURRENT APPLICATION NUMBER: US/09/470,881

CURRENT FILING DATE: 1999-12-22

PRIOR APPLICATION NUMBER: PCT/US99/11780

PRIOR FILING DATE: 1999-05-28

PRIOR APPLICATION NUMBER: 60/087,220

PRIOR FILING DATE: 1998-05-29

NUMBER OF SEQ ID NOS: 8

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 7

LENGTH: 4517

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (208)..(1836)

OTHER INFORMATION: human Yes-1 cDNA translated protein

US-09-470-881-7

Alignment Scores:

Pred. No.: 2,05e-23 Length: 4517

Score: 315.50 Matches: 92

Percent Similarity: 43.22% Conservative: 45

Best Local Similarity: 29.02% Mismatches: 113

Query Match: 23.32% Indels: 67

DB: 4 Gaps: 9

US-09-939-853A-75 (1-261) x US-09-470-881-7 (1-4517)

Qy 2 GlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal--- 20

Db 304 GGAGCAGACCCACTACAGTGTCCATGTCCGTCTTTCAGCAAGGAGGAGCAGCAGTT 363

Qy 21 -----GlnGlyGlnGlyProValThrMetGluAla 30

Db 364 AATTTCAGAGCTTTTCCATGACACCATTTGGAGATCTTCAGGGGTAAACCCCTTTTGA 423

Qy 31 GluArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGly----- 46

Db 424 GGTGATCTTCTCATTTTCAGTGTGCGCAAGTTTCATATCTCTGCTGTTTAAACAGTGGT 483  
Qy 47 -----GlyProAlaGluLeuSerLeu 53  
Db 484 GTTACTATATTGTGGCTTATATGATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 543  
Qy 54 ArgLeuGlyGluProLeuThrIleValSerGlu---AspGlyAspTyrThrValLeu 72  
Db 544 AAGAAGGTTGAAGATTTCAAAATTAATCAATACGAAGAGAGATTTGGTGGGAACAAGA 603  
Qy 73 SerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal----- 90  
Db 604 TCAATCGCTACAGGAAGAAGTGTATATCCGAGCAATATGATGCGCTGAGATTC 663  
Qy 91 -----SerHisGlyTyrLeuTyrGluGlyLeuSerArgGluLysAlaGluGluLeu 108  
Db 664 ATTCAGCGAGAAGATGTTATTTGGCAAAATGGGGAAGAAGATGCTGGAAGATTACT 723  
Qy 109 LeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgGly 128  
Db 724 TTGAATCTGGAAATCAACAGGATTTTCTTAGTAGAGAGAGTGAACACTAAAGT 783  
Qy 129 SerTyrSerLeuSerValArgLeuSerArgProAlaSerTyrAspArgIleArg----- 146  
Db 784 GCTTATTCCTTTCTATTCTGT-----GATTGGATGAGTAAAGGGTGCAC 828  
Qy 147 -----HisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIleSerProArg 163  
Db 829 AATGTGAACACTACAAATATAGGAACCTTGAATGGTGGATACTATATCAACACAGA 888  
Qy 164 LeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspIle 183  
Db 889 GCACAATTTGATCTCTGCAGAAATTTGGTGAACACTACACAGAACATGCTGATGGTTA 948  
Qy 184 CysCysLeuLeuLysGluProCys-----ValLeuGln 194  
Db 949 TCCACAGATTGCAACTGTGTGTCCTCAACTGTGAACCTCAGACTCAAGGTCGTAGCAAA 1008  
Qy 195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArg----- 212  
Db 1009 GATGCTGGGAATCCCTCGAAGTCTTGGCACTAGAGGTTAAACTAGGACAGGATGT 1068  
Qy 213 -----ThrProLeuAsnTyrLysGluLeu 220  
Db 1069 TTCGCGAAGTGTGATGGGAACATGGAATGGAACTGCAAGTGAAGTGAAGTGAAGTGA 1128  
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
Db 1129 AAACAGGTACATGATCCAGAACTTTCTTCAAGAGCTCAGATAATGAAAAATTA 1188  
Qy 241 LeuArgGluSerLeu---SerPheTyrIleSerLeuAsnAspGluAlaVal 256  
Db 1189 AGACATGATAAACTTTGTTCCACTATATGCTGTGTTTCTGAAGAACCAATT 1239

## RESULT 8

PCT-US93-06251-83  
; Sequence 83, Application PC/TUS9306251  
; GENERAL INFORMATION:  
; APPLICANT: Wickstrom, Eric and Rife, Jason P.  
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/06251  
; FILING DATE: 19930630  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8586  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 83:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4517 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; PCT-US93-06251-83

Alignment Scores:  
Pred. No.: 2,05e-23 Length: 4517  
Score: 315.50 Matches: 92  
Percent Similarity: 43.22% Conservative: 45  
Best Local Similarity: 29.02% Mismatches: 113  
Query Match: 23.32% Indels: 67  
DB: 9 Gaps: 9

US-09-939-853A-75 (1-261) x PCT-US93-06251-83 (1-4517)

Qy 2 GlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal---- 20  
Db 304 GGAGCAGAACCCACATACAGTGCACCATGTCGTCATCTTCAGCAAGGGAACAGCAGTT 363  
Qy 21 -----GlnGlyGlnGlyProValThrMetGluAla 30  
Db 364 AATTTCAGCAGTCTTTCATGACACCATTTGGAGGATCCTCAGGGGTAAACGCTTTTGA 423  
Qy 31 GluArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGly----- 46  
Db 424 GGTGCATCTCTCTATTTTCAGTGGTCCCAAGTTTCATATCTCTGCTGTTTAAACAGTGT 483  
Qy 47 -----GlyProAlaGluLeuSerLeu 53  
Db 484 GTTACTATATTTGTGGCTTATATGATGATGAAGCTAGAACCTACAGAACCTTTCAATT 543  
Qy 54 ArgLeuGlyGluProLeuThrIleValSerGlu---AspGlyAspTyrThrValLeu 72  
Db 544 AAGAAGGTTGAAGATTTCAAAATTAATCAATACGAAGAGAGATTTGGTGGGAACAAGA 603  
Qy 73 SerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal----- 90  
Db 604 TCAATCGCTACAGGAAGAAGTGTATATCCGAGCAATATGATGCGCTGAGATTC 663  
Qy 91 -----SerHisGlyTyrLeuTyrGluGlyLeuSerArgGluLysAlaGluGluLeu 108  
Db 664 ATTCAGCGAGAAGATGTTATTTGGCAAAATGGGGAAGAAGATGCTGGAAGATTACT 723  
Qy 109 LeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgGly 128  
Db 724 TTGAATCTGGAAATCAACAGGATTTTCTTAGTAGAGAGAGTGAACACTAAAGT 783  
Qy 129 SerTyrSerLeuSerValArgLeuSerArgProAlaSerTyrAspArgIleArg----- 146  
Db 784 GCTTATTCCTTTCTATTCTGT-----GATTGGATGAGTAAAGGGTGCAC 828  
Qy 147 -----HisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIleSerProArg 163  
Db 829 AATGTGAACACTACAAATATAGGAACCTTGAATGGTGGATACTATATCAACACAGA 888  
Qy 164 LeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspIle 183







QY 38 -----ValAlaLeuGlySerPheProAlaGlyGlyProAlaGlu 50  
 Db 373 GGGATTGGGTGACCCCTGTTCAATGGCCCTGTATGACTATGAGGCTGGAAGTGGATGAC 432  
 QY 51 LeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu---AspGlyAspTTP 69  
 Db 433 CTCACCTTCACCAAGGGCGAAGTCTCCATCTCTGAACAATACTGAAGTGACTGGTGG 492  
 QY 70 ThrValLeuSerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLys 89  
 Db 493 GAGCTCGGTCTCTCAGCTCGGAAACAACTGGCTGCATTCACGCAACTACGTCGGCCCT 552  
 QY 90 Val-----SerHisGlyTTPLeuTyrGluGlyLeuSerArgGluLysAlaGlu 105  
 Db 553 GTTCACTCAATCCAAAGTGAAGTGTACTTTGGAAGATTGGGAAAGGATGCAGAG 612  
 QY 106 GluLeuLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuLeuArgGluSerGlnThr 125  
 Db 613 AGGCAGCTGCTTTCACAGGCAACCCCGGGGGCTTTCTCATTCGGGAAACCGAGACC 672  
 QY 126 ArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIle 145  
 Db 673 ACCAAAGGTGCTACTCCCTGTCATCCGGAGTGGATCAGACAGAGCGATCATGTG 732  
 QY 146 ArgHisTyrArgIleHisCysLeuAspAsnGlyTTPLeuTyrIleSerProArgLeuThr 165  
 Db 733 AAGCATTTACAGATCCGCAAACTGGACATGGGGGGCTTACTATCATCCACACCGGGTTCAG 792  
 QY 166 PheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCys 185  
 Db 793 TTCACCTGGTGGAGGAGTGGTGACCATATATGAGGTGAATCAGCGGGCTGTGCAAC 852  
 QY 186 LeuLeuLysGluProCysValLeuGlnArg 195  
 Db 853 CTGCTCATCGCGCCTGCACCATCATGAAG 882

## RESULT 12

US-09-470-881-2  
 ; Sequence 2, Application US/09470881  
 ; Patent No. 6685938  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHERESH, David A.  
 ; APPLICANT: ELICEIRI, Brian  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR MODULATION OF  
 ; TITLE OF INVENTION: ANGIOGENESIS AND VASCULAR PERMEABILITY USING SRC OR  
 ; TITLE OF INVENTION: YES TYROSINE KINASES  
 ; FILE REFERENCE: TSRI 651.2  
 ; CURRENT APPLICATION NUMBER: US/09/470,881  
 ; PRIOR FILING DATE: 1999-12-22  
 ; PRIOR APPLICATION NUMBER: PCT/US99/11780  
 ; PRIOR FILING DATE: 1999-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087,220  
 ; PRIOR FILING DATE: 1998-05-29  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1759  
 ; TYPE: DNA  
 ; ORGANISM: Chicken  
 ; FEATURE:  
 ; NAME/KEY: gene  
 ; LOCATION: (1)..(1759)  
 ; OTHER INFORMATION: chicken c-SRC cDNA  
 ; NAME/KEY: CDS  
 ; LOCATION: (112)..(1710)  
 US-09-470-881-2

## Alignment Scores:

Pred. No.: 3,35e-21 Length: 1759  
 Score: 289.00 Matches: 86  
 Percent Similarity: 47.33% Conservative: 47  
 Best Local Similarity: 30.60% Mismatches: 104  
 Query Match: 21.36% Indels: 45

DB: 4 Gaps: 6  
 US-09-939-853A-75 (1-261) x US-09-470-881-2 (1-1759)  
 QY 5 ProSerArgArgLysSerLeuProSerProSerLeuSerSerSerValGlnGlyGlnGly 24  
 Db 317 CGTTTACCTGCGCGCAGCGTGGCGGGCAGTGGCTGGCGGGTCAAC----- 363  
 QY 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44  
 Db 364 -----ACTTTCGTGGCTCTCTACGACTACGAG 390  
 QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu 64  
 Db 391 TCCCGGACTGAAACGCGACTTGTCTTCAAGAAAGAGAGAACGCTGCAGATTGTCAACAC 450  
 QY 65 ---AspGlyAspTTPTrpThrValLeuSerGluValSerGlyArgGluTyrAsnIlePro 83  
 Db 451 ACGGAAGGTGACTGGTGGCTGCTCATTCCTCTACTACAGGACACACGGGCTATACCC 510  
 QY 84 SerValHisValGlyLys-----ValSerHisGlyTTPLeuTyrGluGlyLeu 99  
 Db 511 AGTAACACTATGTCGCCCTCAGACTCCATCCAGGCTGAAGAGTGTACTTTGGGAAGATC 570  
 QY 100 SerArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeu 119  
 Db 571 ACTCGTCGGAGTCGAGCGGCTGCTCAACCCCGGAAACCCCGGGGAAACCTTCTTG 630  
 QY 120 IleArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgPro 139  
 Db 631 GTCCGGGAGAGCGAGACAGACAAAGGTGCTATTGCTCTCCGTTCTTGACTTTGACAAAC 690  
 QY 140 AlaSerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTTPLeuTyr 159  
 Db 691 GCCAAGGGGCTCAATGTGAGCAGCTACAAGATCCGCAAGCTGGACAGCGGGCTTCTAC 750  
 QY 160 IleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeu 179  
 Db 751 ATCACTCAGCAGACACAGTTCAGCAGCTGCGAGAGCTGGTGGCTTACTTCTCAAAACAT 810  
 QY 180 AlaAspAspIleCysCysLeuLysGluProCys-----ThrProLeuAsn 216  
 Db 811 GCTGATGCTTGTGCCCGCGCTGACCAACGCTGCCCCAGCTCAAGCCGCCAGCCAG 870  
 QY 192 ---ValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrVal 210  
 Db 871 GGACTCGCAAGGACGCGTGGGAAATCCCGGGAGTGGCTGCGGCTGGAGGTGAAGCTG 930  
 QY 211 GlnArg-----ThrProLeuAsn 216  
 Db 931 GGGCAGGCTCTTTGGAGAGTCTGGATGGGGACCTGGAAACGGCACACAGAGTGGCC 990  
 QY 217 TrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeu 236  
 Db 991 ATAAAGACTCTGAAGCCCGGACCATGTCCCGGAGGCTTCTCTGCAGGAAGCCCAAGTG 1050  
 QY 237 LeuSerGluGlyLeuArgGluSerLeu---SerPheTyrIleSerLeuAsnAspGluAla 255  
 Db 1051 ATGAAGAAGTCCCGGCATGAGAAGCTGTTGAGCTGACGAGTGGTGGAGAGAGGCC 1110  
 QY 256 Val 256  
 Db 1111 ATC 1113

## RESULT 13

US-07-820-011A-1  
 ; Sequence 1, Application US/07820011A  
 ; Patent No. 5336615  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell, Leonard  
 ; APPLICANT: Madri, Joseph A.  
 ; APPLICANT: Warren, Stephen L.  
 ; APPLICANT: Luthringer, Daniel J.

TITLE OF INVENTION: Genetically Engineered  
TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced  
TITLE OF INVENTION: Migration  
TITLE OF INVENTION: And Plasminogen Activator Activity  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb storage  
COMPUTER: IBM PC XT  
OPERATING SYSTEM: PC-DOS/MS-DOS 2.10  
SOFTWARE: Displaywrite 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820.011A  
FILING DATE: 19920106  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: LB-101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255 1400  
TELEFAX: (203) 254 1101  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1602 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: No  
ANTI-SENSE: No

ORIGINAL SOURCE:  
ORGANISM: Gallus, gallus  
PUBLICATION INFORMATION:  
AUTHORS: Tanaka, Tatsuo  
AUTHORS: Hanafusa, Hideaburo  
TITLE: Structure and Sequence of the RSV src  
TITLE: Cellular Gene Homologous to the RSV src  
TITLE: Gene and the Mechanism for Generating the  
TITLE: Transforming Virus  
JOURNAL: Cell  
VOLUME: 32  
PAGES: 881-890  
DATE: March, 1983

US-07-820-011A-1  
Alignment Scores:  
Pred. No.: Length: 4.75e-21 1602  
Score: 287.00 Matches: 81  
Percent Similarity: 51.28% Conservative: 39  
Best Local Similarity: 34.62% Mismatches: 84  
Query Match: 21.21% Indels: 32  
DB: 1 Gaps: 5

US-09-939-853a-75 (1-261) x US-07-820-011A-1 (1-1602)

QY 5 ProSerArgAlaGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24

Db 206 CCGTTACGTCGCCGAGCGTCCCGGGACGTCGGTGGCGGCTCACCC----- 252

QY 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44

Db 253 -----ACTTTGCTGGCTCTCTACGACTACGAG 279

QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu 64

Db 280 TCCCGAGCTGAACCGGACTTGTCTTCAAGAAAGAGAACCGCTGCAGATTGTCAACAC 339

65 ---AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTrpAsnIlePro 83

Db 340 ACGGAAGGTGACTGGTGGCTCATTCCTCACTACAGGACAGCGGGCTACATCCCC 399

QY 84 SerValHisValGlyLys-----ValSerHisGlyTrpLeuTrpGluGlyLeu 99

Db 400 AGTAACATATGTCGGCCCTCAGACTCCATCCAGGCTGAAGAGTGGTACTTTGGGAAGATC 459

QY 100 SerArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeu 119

Db 460 ACTCGTCGGGAGTCGAGCGGCTGCTCAACCCCGAAACCCCGGGGAACCTTCTTG 519

QY 120 IleArgGluSerGlnThrArgArgGlySerTrpSerLeuSerValArgLeuSerArgPro 139

Db 520 GTCCGGGAGAGCGAGACGACAAAGGTCCTATTCCTCTCCGTTCTCGACTTTGACAAAC 579

QY 140 AlaSerTrpAspArgIleArgHisTrpArgIleHisCysLeuAspAsnGlyTrpLeuTrp 159

Db 580 GCCAAGGGGCTCAATGTGAGCACTACAGATCCGCAAGCTGGACAGCGGGCTTCTAC 639

QY 160 IleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAlaPheHisTrpSerGluLeu 179

Db 640 ATCACTCAGCAGACACAGTTCAGCAGCTGCAGAGCTGGTGGCTACTACTCCAAACAT 699

QY 180 AlaAspAspIleCysCysLeuLeuLysGluProCys----- 191

Db 700 GCTGATGCTTGTGCCACCGCTGACCAACGCTGCCCCACGTCACAGCCCGACCCAG 759

QY 192 ---ValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrVal 210

Db 760 GGACTCGCCAAAGACGCGTGGGAAATCCCGGGAGTGGCTGGCGCTGGAGGTGAAGCTG 819

QY 211 GlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224

Db 820 GGGCAGGCGCTG-CIT---TGGAGAGCTGTGATGGGGACCTG 857

RESULT 14

PCT-US93-00445-1

; Sequence 1, Application PC/TUS9300445

; GENERAL INFORMATION:

; APPLICANT: Bell, Leonard

; APPLICANT: Madri, Joseph A.

; APPLICANT: Warren, Stephen L.

; APPLICANT: Luthringer, Daniel J.

; TITLE OF INVENTION: Genetically Engineered

; TITLE OF INVENTION: Endothelial Cells

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maurice M. Klee

; STREET: 1951 Burr Street

; CITY: Fairfield

; STATE: Connecticut

; COUNTRY: USA

; ZIP: 06430

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 760 Kb storage

; COMPUTER: DELL 486/50

; OPERATING SYSTEM: DOS 5.0

; SOFTWARE: Displaywrite 3

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/00445

; FILING DATE: 19930105

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/820,011

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Klee, Maurice M.

; REGISTRATION NUMBER: 30,399

; REFERENCE/DOCKET NUMBER: ALX-101PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (203) 255 1400

```

; TELEFAX: (203) 254 1101
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1602 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: Gallus, gallus
; PUBLICATION INFORMATION:
; AUTHORS: Takeya, Tatsuo
; AUTHORS: Hanafusa, Hidesaburo
; TITLE: Structure and Sequence of the
; TITLE: Cellular Gene Homologous to the RSV src
; TITLE: Gene and the Mechanism for Generating the
; TITLE: Transforming Virus
; JOURNAL: Cell
; PAGES: 32
; VOLUME: 81-890
; DATE: March, 1983
; PCT-US93-00445-1

Alignment Scores:
Pred. No.: 4,75e-21 Length: 1602
Score: 287.00 Matches: 81
Percent Similarity: 51.28% Conservative: 39
Best Local Similarity: 34.62% Mismatches: 84
Query Match: 21.21% Indels: 32
DB: 5 Gaps: 5

US-09-939-853a-75 (1-261) x PCT-US93-00445-1 (1-1602)

QY 5 ProSerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24
Db 206 CGTTAGTCTGGCGAGCGTGGCGGCGACTGCTGGCGGCGTACAC----- 252
QY 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44
Db 253 -----ACTTTGTGGCTCTACGACTACGAG 279
QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGlyProLeuThrIleValSerGlu 64
Db 280 TCCGGGACTGAACGAGCTTGCTTCAAGAAGGAGAACGCTGCAGATTGTCAACAAC 339
QY 65 ---AspGlyAspTrpThrValLeuSerGluValSerGlyValSerGlyGlyTrpAsnIlePro 83
Db 340 ACGGAAGGTGACTGTGGTGGTGCATTCCTCCTCACTACAGGACAGCGGCTACATCCCC 399
QY 84 SerValHisValGlyLys-----ValSerHisGlyTrpLeuTyrrGluGlyLeu 99
Db 400 AGTAACATATGTCGGCGCCTCAGACTCATCCAGGTGAAGAGGTGACTTTGGGAAGATC 459
QY 100 SerArgGlyLysAlaGluGluLeuLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeu 119
Db 460 ACTCGTCGGGAGTCCGAGCGGCTGCTGCTCAACCCCGAAAAACCCCGGGGAACTTCTTG 519
QY 120 IleArgGluSerGlnThrArgArgGlySerTyrrSerLeuSerValArgLeuSerArgPro 139
Db 520 GTCCGGGAGAGCGAGCAGACAAAGGTGCTATGTGCTCTCTCGCTTCTGACTTTGACACAC 579
QY 140 AlaSerTrpAspArgIleArgHisTrpArgIleHisLeuAspAsnGlyTrpLeuTyrr 159
Db 580 GCCAAGGGGCTCAATGTGAAGGACTACAGATCCGCAAGCTGGACAGCGGCGGCTTCTAC 639
QY 160 IleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrrSerGluLeu 179
Db 640 ATCACTTCAGCACACAGATTTCAGCAGCTGCAGCAGCTGTGCTGCTACTACTCCAAACAT 699
QY 180 AlaAspAspIleCysCysLeuLeuLysGluProCys----- 191

```

700 GCTGATGCTTGTGCCACCGCTGACCAACGCTGTGCCCCACGTCACAGCCCGACCCAG 759  
 192 ---ValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrVal 210  
 760 GGACTCGCAAGAGGAGCGTGGGAAATCCCGGAGTGTGCTGCGCTGGAGGTGAAGCTG 819  
 211 GlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224  
 820 GGGCAGGCGTG-CTT---TGGAGAGGTCTGGATGGGACCTG 857

RESULT 15  
 US-07-820-011A-3  
 ; Sequence 3, Application: US/07820011A  
 ; Patent No. 5336615  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell, Leonard  
 ; APPLICANT: Madri, Joseph A.  
 ; APPLICANT: Warren, Stephen L.  
 ; APPLICANT: Iuthringer, Daniel J.  
 ; TITLE OF INVENTION: Genetically Engineered  
 ; TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced  
 ; TITLE OF INVENTION: Migration  
 ; TITLE OF INVENTION: and Plasminogen Activator Activity  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Maurice M. Klee  
 ; STREET: 1951 Burr Street  
 ; CITY: Fairfield  
 ; STATE: Connecticut  
 ; COUNTRY: USA  
 ; ZIP: 06430  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb storage  
 ; COMPUTER: IBM PC XT  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS 2.10  
 ; SOFTWARE: Displaywrite 3  
 ; CURRENT APPLICATION DATA:  
 ; FILING DATE: 19920106  
 ; APPLICATION NUMBER: US/07820,011A  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Klee, Maurice M.  
 ; REGISTRATION NUMBER: 30,399  
 ; REFERENCE/DOCKET NUMBER: LB-101  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (203) 255 1400  
 ; TELEFAX: (203) 254 1101  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1611  
 ; TYPE: NUCLEIC ACID  
 ; STRANDEDNESS: Double  
 ; TOPOLOGY: Linear  
 ; MOLECULE TYPE: cDNA to mRNA  
 ; HYPOTHETICAL: No  
 ; ANTI-SENSE: No  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapien  
 ; POSITION IN GENOME:  
 ; CHROMOSOME/SEGMENT: Chromosome 20  
 ; PUBLICATION INFORMATION:  
 ; AUTHORS: Anderson, Stephen K.  
 ; AUTHORS: Gibbs, Carol P.  
 ; AUTHORS: Tanaka, Akio  
 ; AUTHORS: Kung, Hsing-Jien  
 ; AUTHORS: Fujita, Donald J.  
 ; TITLE: Human Cellular src Gene:  
 ; TITLE: Nucleotide Sequence and Derived Amino  
 ; TITLE: Acid Sequence of the Region Coding for  
 ; TITLE: the Carboxy-Terminal Two-Thirds of  
 ; TITLE: pp60c-src  
 ; JOURNAL: Molecular and Cellular Biology  
 ; VOLUME: 5

ISSUE: 5  
PAGES: 1122-1129  
DATE: May, 1985  
PUBLICATION INFORMATION:  
AUTHORS: Tanaka, Akio  
AUTHORS: Gibbs, Carol P.  
AUTHORS: Arthur, Richard R.  
AUTHORS: Anderson, Stephen K.  
AUTHORS: Kung, Hsing-Jien  
AUTHORS: Fujita, Donald J.  
TITLE: DNA Sequence Encoding the  
TITLE: Amino-Terminal Region of the Human c-src  
TITLE: Protein: Implications of Sequence  
TITLE: Divergence among src-Type Kinase  
TITLE: Oncogenes  
JOURNAL: Molecular and Cellular Biology  
VOLUME: 7  
PAGES: 1978-1983  
ISSUE: 5  
DATE: May, 1987  
US-07-820-011A-3

Alignment Scores:  
Pred. No.: 5.41e-20 Length: 1611  
Score: 276.50 Matches: 87  
Percent Similarity: 46.15% Conservative: 45  
Best Local Similarity: 30.42% Mismatches: 117  
Query Match: 20.44% Indels: 38  
DB: 1 Gaps: 6

US-09-939-853a-75 (1-261) x US-07-820-011A-3 (1-1611)

QY 8 ArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGlnGlyProValThr 27  
DB 155 CGGCCCTTCGCCCGCGCGCCGCGCCGAGCCCAAGCTGTCGGAGGCTTCAACTCTCTCGGACA 214  
QY 28 MetGluAlaGluArgSerLysAla-----ThrAlaValAla 39  
DB 215 CGGTCACTCCCGCGCAG-AGGCGCGCGCGCGTGGCGGTGAGTGACCACTTTGTGGCC 273  
QY 40 LeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeu 59  
DB 274 CTCATGACTATGACTAGACGGAGACACACTGTCTTCAAGAAGCGGAGCGGCTC 333  
QY 60 ThrIleValSerGlu---AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArg 78  
DB 334 CAGATTGTCAACACACAGAGGAGAGCTGTGGTGGCCCACTCGCTCAGCACAGGACAG 393  
QY 79 GluTyrAsnIleProSerValHisValGlyLys-----ValSerHisGlyTrp 94  
DB 394 ACAGCTACATCCCGCAGCACTAGCTGGCGCCCTCCGACTCCATCCAGGCTGAGGAGTGG 453  
QY 95 LeuTyrGluGlyLeuSerArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnPro 114  
DB 454 TATTTGGCAAGATCACCACAGCGGAGTCAGAGCGTTACTGTCTCAATGCAGAGAACCCG 513  
QY 115 GlyGlyAlaPheLeuIleArgGluSerGlnThrArgGlySerTyrSerLeuSerVal 134  
DB 514 AGAGGGACCTTCCTCGTCGGAAGAGTGAGCACCAAGAGTGGCTACTGCTCTCAGTG 573  
QY 135 ArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIleHisCysLeuAsp 154  
DB 574 TCTGACTTCGACAAAGCGGCTCAACGCTGAGCAGCTACAGATCCGCAAGCTGGAC 633  
QY 155 AsnGlyTrpLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAsp 174  
DB 634 AGCGCGCGCTTCTACATCCTCCCGCACCAGTTCACAGCCCTGCAGCAGCTGGTGGCC 693  
QY 175 HisTyrSerGluLeuAlaAspIleCysCysLeuLeuLysGluProCys----- 191  
DB 694 TACTACTCAACACAGCGGAGTGGCTGTGCCACCGCTCACCACCGGTGTGCCACCGTCC 753  
QY 192 -----ValLeuGlnArgAlaGlyProLeuProGlyLysAspIlePro 205

Search completed: February 20, 2004, 11:33:00  
Job time : 120 secs

DB 754 AAGCCGCACTCAGGGCTGGCCCAAGGATCCCTGGGAGATCCCTCGGAGTGGCTGCGG 813  
QY 206 LeuProValThrValGlnArg-----  
DB 814 CTGGAGGTCAAGCTGGCCAGGCTGCTTTGGCGAGGTGTGATGGGACCTGGAAACGGT 873  
QY 213 ----ThrProLeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThr 231  
DB 874 ACCACCAGGCTGGCCATCAAAACCCCTGAAGCTGGCACGATGTCTCCAGAGGCTTCTG 933  
QY 232 GlyGluGluSerLeuLeuSerGluGlyLeuArgGluSerLeu---SerPheTyrIleSer 250  
DB 934 CAGGAGGCCAGGTCTATGAGAGCTGAGGATGAGAAAGCTGGTGCAGTTGTATGCTGTG 993  
QY 251 LeuAsnAspGluAlaVal 256  
DB 994 GTTTCAGAGGAGCCCAATT 1011

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 10:48:39 ; Search time 357 Seconds  
(without alignments)  
2560.026 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSUPSRKSLPSPSSSV.....RESLSFYISLNDVAVSLDDA 261

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-O=/cn2\_1/USPTO spool\_p/US0939853/runat\_19022004\_145341\_24590/app\_query.fasta\_1.455  
-D=Published Applications\_NA\_QMWT-fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -MAP=US0939853 @CGN 1 1 333 @runat 19022004 145341\_24590  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cn2\_6/ptodata/2/pubpna/PTC\_NEW\_PUB.seq.\*
- 3: /cn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cn2\_6/ptodata/2/pubpna/PTC\_US\_PUBCOMB.seq.\*
- 7: /cn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	1347	99.6	786	14	US-10-043-649-1	Sequence 1, Appli
2	826	61.0	783	9	US-09-867-550-953	Sequence 953, App
3	643	47.5	864	10	US-09-814-353-21302	Sequence 21302, A
4	586	43.3	875	9	US-09-867-550-1915	Sequence 1915, App
5	488	36.1	3756	13	US-10-002-600-91	Sequence 91, Appl
6	487	36.0	2655	9	US-09-954-456-499	Sequence 499, App
7	452.5	33.4	444	9	US-09-867-550-951	Sequence 951, App
8	370.5	27.4	2298	14	US-10-175-523-50	Sequence 50, Appl
9	370.5	27.4	2298	15	US-10-153-563-343	Sequence 343, App
10	360.5	26.6	1924	15	US-10-193-720-1	Sequence 1, Appli
11	360.5	26.6	2015	9	US-09-954-456-1983	Sequence 1983, Ap
12	360.5	26.6	2015	14	US-10-007-010-3	Sequence 3, Appli
13	360.5	26.6	2341	14	US-10-252-157-140	Sequence 140, App
14	360.5	26.6	2341	15	US-10-062-674-2038	Sequence 2038, Ap
15	358.5	26.5	1911	9	US-09-917-8008-1611	Sequence 1611, Ap
16	350	25.9	320	10	US-09-814-353-17314	Sequence 17314, A
17	348	25.7	2032	15	US-10-366-288-27	Sequence 27, Appl
18	340	25.1	2017	15	US-10-062-674-1776	Sequence 1776, Ap
19	340	25.1	2129	15	US-09-960-706-954	Sequence 954, App
20	340	25.1	2129	15	US-10-305-720-1452	Sequence 1452, Ap
21	324	23.9	2433	14	US-10-240-965-114	Sequence 114, App
22	322.5	23.8	2451	9	US-09-771-161A-4	Sequence 4, Appli
23	320	23.7	1609	9	US-09-771-161A-30	Sequence 30, Appl
24	320	23.7	1995	9	US-09-771-161A-31	Sequence 31, Appl
25	320	23.7	4414	14	US-10-101-510-512	Sequence 512, App
26	315.5	23.3	4343	15	US-10-159-563-184	Sequence 184, App
27	315.5	23.3	4517	14	US-10-298-377A-3	Sequence 3, Appli
28	315.5	23.3	4517	14	US-10-007-926A-304	Sequence 304, App
29	314.5	23.2	549	10	US-09-918-995-25044	Sequence 25044, A
30	312.5	23.1	1590	15	US-10-085-117-18	Sequence 18, Appl
31	312.5	23.1	2354	9	US-09-967-768A-300	Sequence 300, App
32	312.5	23.1	2354	15	US-10-353-690-123	Sequence 123, App
33	312.5	23.1	2354	15	US-10-085-117-17	Sequence 17, Appl
34	303	22.4	1554	15	US-10-085-117-15	Sequence 15, Appl
35	303	22.4	2179	15	US-10-085-117-14	Sequence 14, Appl
36	293.5	21.0	488	10	US-09-918-995-19745	Sequence 19745, A
37	246	18.2	432	9	US-09-864-761-2829	Sequence 2829, Ap
38	246	18.2	448	9	US-09-864-761-15513	Sequence 15513, A
39	240.5	17.8	2770	9	US-09-977-269-5	Sequence 5, Appli
40	240.5	17.8	2770	10	US-09-977-261-5	Sequence 5, Appli
41	240.5	17.8	2770	10	US-09-977-261-5	Sequence 1631, Ap
42	240.5	17.8	2863	9	US-09-954-456-1631	Sequence 19, Appl
43	240.5	17.8	7607	9	US-09-982-610-19	Sequence 1087, Ap
44	238.5	17.6	2173	15	US-10-094-749-1087	Sequence 1242, Ap
45	230.5	17.0	342	15	US-10-062-674-1242	

## ALIGNMENTS

### RESULT 1

US-10-043-649-1

; Sequence 1, Application US/10043649

; Publication No. US20030059924A1

; GENERAL INFORMATION:

; APPLICANT: Holland, Sacha J.

; APPLICANT: Menderhall, Marcy K.

; APPLICANT: Pardo, Jorge

; APPLICANT: Spencer, Collin

; APPLICANT: Fu, C. Alan

; APPLICANT: Luo, Ying

; APPLICANT: Payan, Donald G.

; APPLICANT: Mancebo, Helena S.Y.

; APPLICANT: Wu, Jun

; APPLICANT: Zhou, Xiulan

; APPLICANT: Shen, Mary

; APPLICANT: Liao, X. Charlene

; APPLICANT: Sheng, Ning

; TITLE OF INVENTION: Cloning of a No. US20030059924A1e1 Inhibitor of Antigen-receptor

; FILE REFERENCE: A-70219-1/RMS/DHR

; CURRENT APPLICATION NUMBER: US/10/043,649

; CURRENT FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 60/260,953

; PRIOR FILING DATE: 2001-01-10  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 786  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1..786)  
 ; OTHER INFORMATION:  
 US-10-043-649-1

Alignment Scores:  
 Pred. No.: 6,31e-146 Length: 786  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.56% Indels: 0  
 DB: 14 Gaps: 0

US-09-939-853A-75 (1-261) x US-10-043-649-1 (1-786)

Qy	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db	1	ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTCCCAAGCCCAAGCTTGAGTTCTCTGTC	60
Qy	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	61	CAGGCCAGGAGCTGTGACCATGGAGCAGAGAGAGAGCCACAGCCGTCGGCCCTG	120
Qy	41	GlySerPheProAlaGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
Db	121	GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTCTGAGACTCGGGAGCCATTGACC	180
Qy	61	IleValSerGluAspGlyAspTrrThrValLeuSerGluValSerGlyArgGluTyr	80
Db	181	ATGCTCTCTGAGGATGAGACTGGTGACCGTGTCTGTAAGTCTCAGCGCAGAGATAT	240
Qy	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrrPleuTyrGluGlyLeuSer	100
Db	241	AACATCCCCAGCTCCACGTGGCCAAAGTCTCCATGGTGGCTGTATGAGGCCCTGAGC	300
Qy	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuLeu	120
Db	301	AGGAGAAAGCAGAGAACTGCTGTGTACCTGGGAACCTTGAGGGCCCTTCTCTCATC	360
Qy	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	361	CGGAGAGCCAGACAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGCA	420
Qy	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrrPleuTyrIle	160
Db	421	TCCTGGACCCGATCAGACACTCAGATCCACTGCTTTGACAATGGCTGGCTGTACATC	480
Qy	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
Db	481	TCACCGGCTCTACCTTCCCTCACTCCAGCCCTGGTGAGACCATTACTCTGAGCTGGCG	540
Qy	181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
Db	541	GATGACATCTGCTGCTCTACTCAAGAGCCCTGTGTCTGTCAGAGGGCTGGCCGCTCCCT	600
Qy	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
Db	601	GGCAGAGATATACCTTACCTGTGACTGTGCAGAGAGACCACTCACTCACTGGAAGAGCTG	660
Qy	221	AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGly	240
Db	661	GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGGCTCTTCTCAGTGAAGGT	720
Qy	241	LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp	260

Db	721	CTCCGGGAGTCCCTCAGCTTACATCAGCGCTGAATGACGAGGCTGTCTCTTTTGGATGAT	780
Qy	261	Ala 261	
Db	781	GCC 783	

RESULT 2  
 US-09-867-550-953  
 ; Sequence 953, Application US/09867550  
 ; Patent No. US20020982206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad,  
 ; APPLICANT: Conley, Pamela,  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20020982206A1 Polynucleotides from Atherogenic Cells and  
 ; FILE REFERENCE: 21402-013 (Cura-313)  
 ; CURRENT APPLICATION NUMBER: US/09/867,550  
 ; CURRENT FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: USN 60/208,427  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 953  
 ; LENGTH: 763  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-867-550-953

Alignment Scores:  
 Pred. No.: 9,35e-86 Length: 763  
 Score: 826.00 Matches: 158  
 Percent Similarity: 99.37% Conservative: 0  
 Best Local Similarity: 99.37% Mismatches: 1  
 Query Match: 61.05% Indels: 0  
 DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x US-09-867-550-953 (1-763)

Qy	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db	286	ATGGGAAGTCTGCCAGCAGAGAAATCTCTCCCAAGCCCAAGCTTGAGTTCTCTGTC	345
Qy	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	346	CAAGGCCAGGAGCTGTGACCATGGAGCAGAGAGAGAGCCACAGCCGTCGCCCTG	405
Qy	41	GlySerPheProAlaGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
Db	406	GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTCTGAGACTCGGGAGCCATTGACC	465
Qy	61	IleValSerGluAspGlyAspTrrThrValLeuSerGluValSerGlyArgGluTyr	80
Db	466	ATCCTCTCTGAGGATGAGACTGGTGACCGTGTCTGTAAGTCTCAGCGCAGAGATAT	525
Qy	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrrPleuTyrGluGlyLeuSer	100
Db	526	AACATCCCCAGCTCCACGTGGCCAAAGTCTCCATGGTGGCTGTATGAGGCCCTGAGC	585
Qy	101	ArgGluLysAlaGluGluLeuLeuProGlyAsnProGlyAlaPheLeuLeu	120
Db	586	AGGAGAAAGCAGAGAACTGCTGTGTACTTACCTGGGAGCCCTGGAGGGGCTTCTCTATC	645
Qy	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	646	CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTGTCTGTCGCTCAGCGCCCTGCA	705
Qy	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrrPleuTyr	159
Db	706	TCCTGGGAGCCGATCAGACACTACAGATCCACTGCTTGCACATGCTGGCTGGCTGAC	762

## RESULT 3

US-09-814-353-21302  
; Sequence 21302, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21302  
; LENGTH: 864  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1, 2, 3, 32, 862, 863, 864  
; OTHER INFORMATION: n = A,T,C or G

## US-09-814-353-21302

Alignment Scores:  
Pred. No.: 1,54e-64 Length: 864  
Score: 643.00 Matches: 126  
Percent Similarity: 98.45%  
Best Local Similarity: 97.67% Mismatches: 2  
Query Match: 47.52% Indels: 0  
DB: 10 Gaps: 0

## US-09-939-853A-75 (1-261) x US-09-814-353-21302 (1-864)

QY	1	MetGlySerLeuProSerArgA:GlySerLeuProSerProSerLeuSerSerVal	20
Db	450	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGATTCCTCTGTC	509
QY	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	510	CAAGGCCAGGGACTGTGACCATGGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG	569
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLenglyGluProLeuThr	60
Db	570	GGCAGTTCCCGCAGGTGGCCCGCGAGCTCTCTGAGACTCGGGAGCCATTGACC	629
QY	61	IleValSerGluAspGlyAspThrThrValLeuSerGluValSerGlyArgGluTyr	80
Db	630	ATCGTCTCTGAGGATGAGACTGGTGACGGTCTCTGTAAGTCTTCAGGCAGAGAGTAT	689
QY	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
Db	690	AACATCCCAAGCTCCACGTCGCAAGTCTCCATGGGTGGCTGTATGAGGCCCTGAGC	749
QY	101	ArgGluLysAlaGluLeuLeuLeuLeuLeuLeuProGlyAsnProGlyAlaPheLeuIle	120
Db	750	AGGGAGAAACAGAGAACTGCTGTTGTACCTGGGAACCCCTGGAGGGCCCTTCCTCATC	809
QY	121	ArgGluSerGlnThrArgArgGlySer	129

## RESULT 4

US-09-867-550-1915  
; Sequence 1915, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and i  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1915  
; LENGTH: 875  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)\_feature  
; OTHER INFORMATION: Wherein n is one of a or t or c or g

## US-09-867-550-1915

Alignment Scores:  
Pred. No.: 6,02e-58 Length: 875  
Score: 586.00 Matches: 112  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 43.31% Indels: 0  
DB: 9 Gaps: 0

## US-09-939-853A-75 (1-261) x US-09-867-550-1915 (1-875)

QY	150	IleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeuThrPheProSerLeu	169
Db	4	ATCCACTGCCCTTGACAAATGGCTGTACATCTCACCGCGCCTCAGCTTCCCTCACTC	63
QY	170	GlnAlaLeuValAspHisTyrSerGluLeuAlaAspIleCysCysLeuLeuLysGlu	189
Db	64	CAGGCCCTGGTGGACCATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTCAAGGAG	123
QY	190	ProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThr	209
Db	124	CCCTGTGTCTGTCAGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACT	183
QY	210	ValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAla	229
Db	184	GTGCAGAGGACACCACTCACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAGACT	243
QY	230	AlaThrGlyGluSerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIle	249
Db	244	GCCACAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATC	303
QY	250	SerLeuAsnAspGluAlaValSerLeuAspAla	261
Db	304	AGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCC	339

## RESULT 5

US-10-002-600-91  
; Sequence 91, Application US/10002600  
; Publication No. US20020137077A1  
; GENERAL INFORMATION:  
; APPLICANT: Hopkins, Christopher M.  
; APPLICANT: Peterson, David P.



```

; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS
; FILE REFERENCE: PA-0042 US
; CURRENT APPLICATION NUMBER: US/10/002,600
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 60/243,521
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PERL Program
; SEQ ID NO 91
; LENGTH: 3756
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Template ID: 059263.15
US-10-002-600-91

Alignment Scores:
Pred. No.: 9,54e-46 Length: 3756
Score: 488.00 Matches: 101
Percent Similarity: 57.20% Conservative: 46
Best Local Similarity: 39.30% Mismatches: 94
Query Match: 36.07% Indels: 16
DB: 13 Gaps: 3

US-09-939-853A-75 (1-261) x US-10-002-600-91 (1-3756)

Qy 5 ProSerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24
Db 1098 CCAAGGAAAGAAAGAAATGGGAACAGCATGAAATCCACCCCTCGCTCCGAGAGG 1157
Qy 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44
Db 1158 CCCTGCCCCAACCGGAGGACTGGTAGCGACTTCTCCCGTGTAACTGACTACCG 1217
Qy 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGlnProLeuThrLeuValSerGlu 64
Db 1218 TCTCTGACATCAGCCCGCGATATTCGCGGAGGAGAACTGGGTGTGATTTCTGAT 1277
Qy 65 AspGlyAspTrpThrValLeuSerGluValSerGlyValArgGlyLysLeuValSer 84
Db 1278 GAAGGGGCTGGTGAAGAGTATTTCTTAGCAGTGTTCAGAGAGTATCATCTCTGA 1337
Qy 85 ValHisValGlyLysValSerHisGlyTrpLeuTrpGlyGlyLeuSerArgGlyLysAla 104
Db 1338 ATATGTGTCGACAGAGTTTACTCATGCTGCTGTTGAGGCGCTGGGCGAGACAAGGC 1397
Qy 105 GluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeuArgGlySerGln 124
Db 1398 GAGGAGCTGCTGAGCTGCCAGACACAAGGTGGCTCTTCATGATCAGAGAGTGAG 1457
Qy 125 ThrArgArgGlySerTrpSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144
Db 1458 ACCAAGAAGGTTTACTCATCTGCTGCTGAGACAAAGG-----CAG 1499
Qy 145 IleArgHisTrpArgGlyLeuHisCysLeuAspAsnGlyTrpLeuTrpLeuSerProArgLeu 164
Db 1500 GTAAGGATTAACCGCATTTTCGCTGCCCAACACTGGTACTACATATTTCCCGAGGCTC 1559
Qy 165 ThrPheProSerLeuGlnAlaLeuValAlaPheHisTrpSerGluLeuAlaAspAspLeuCys 184
Db 1560 ACCTTCCAGTCCCTGGAGGACTGTTGAACCACTATTCTGAGGTGGCTGATGCCGTGTC 1619
Qy 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspLeu 204
Db 1620 TGTGTGCTCACACGCGCTGCTGACACAAAGCAGCGCTGCCCGAGCAGTGAGCGCTCC 1679
Qy 205 ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224
Db 1680 AGCTCACCTGTACCTTGGCTCAGAGACTGTGGACTGGAGAGAGTGTCCAGA----- 1733

```

```

Qy 225 LeuPheSerGluAlaAlaThrGly-----GluGluSerLeuLeu 237
Db 1734 --CTGAGGAGGAGCCCGGAGGACAGAGAACCCCTTGGGTAGACGAGTCCCTTTTC 1790
Qy 238 SerGluGlyLeuArgGluSerLeuSerPheTrpLeuSerLeuAsnAspGlu 254
Db 1791 AGCTATGGCTTCGAGAGAGCATTGCTCTTACTGCTCCCTGACCAGTGAG 1841

RESULT 6
US-09-954-456-499
; Sequence 499, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 499
; LENGTH: 2665
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-499

Alignment Scores:
Pred. No.: 7,73e-46 Length: 2665
Score: 487.00 Matches: 101
Percent Similarity: 57.20% Conservative: 46
Best Local Similarity: 39.30% Mismatches: 94
Query Match: 35.99% Indels: 16
DB: 9 Gaps: 3

US-09-939-853A-75 (1-261) x US-09-954-456-499 (1-2665)

Qy 5 ProSerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24
Db 24 CAGGGAAGAAAGAAAGAAATGGGAACAGCATGAAATCCACCCCTCGCTCCGAGAGG 83
Qy 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44
Db 84 CCCCCTGCCAACCCCGGAGGACTGGTAGCGACTTCTTGGCGTGTAACTGACTACCG 143
Qy 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGlnProLeuThrLeuValSerGlu 64
Db 144 TCTCTCATCATCAGCCCCCGCATATTCCCGAGGAGGAGAACTGCGGTGATTTCTGAT 203
Qy 65 AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGlyLysLeuValSerPro 84
Db 204 GAAGGGGCTGTGTGAAAGCTATTCTTCTTAGCAGTGTGTCGAGAGAGTTACATCCCTGGA 263

```

QY 85 ValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSerArgGluLysAla 104  
 ...  
 Db 264 ATATGTGGCCAGAGATTACCATGGCTGGCTGTGGGGCCCTGGGAGAGACAAGGCC 323  
 ...  
 QY 105 GluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeuArgGluSerGln 124  
 ...  
 Db 324 GAGGAGCTGCTGACCTGCCACACACAAGGTGGCTCTTCATGATCAGAGAGATGAG 383  
 ...  
 QY 125 ThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144  
 ...  
 Db 384 ACCAAGAAGGGTTTACTCACTGCGGTGAGA-----CACAGGCGAG 425  
 ...  
 QY 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeu 164  
 ...  
 Db 426 GAAAGCATACCGATTTCCTGCTGCCGAACAACCTGTACTACATTTCCCGAGGCTC 485  
 ...  
 QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCys 184  
 ...  
 Db 486 ACCTCCAGTGGCTGGAGGACCTGGTGAACCACTATTCTGAGTGGCTGATGGCTGTGC 545  
 ...  
 QY 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
 ...  
 Db 546 TGTGTGCTCACCAGCCCTGCTGCTGAGAGACTGTGGACTGGAGGAGAGTGTCCAGA----- 659  
 ...  
 QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224  
 ...  
 Db 606 AGCTACCTGTCACTTGGCTGAGAGACTGTGGACTGGAGGAGAGTGTCCAGA----- 659  
 ...  
 QY 225 LeuPheSerGluAlaAlaThrGly-----GluGluSerLeuLeu 237  
 ...  
 Db 660 ---CTGCAGAGAGACCCCGAGGAGAACAGAGAACCCGCTGGGGTAGAGAGTCCCTTTTC 716  
 ...  
 QY 238 SerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu 254  
 ...  
 Db 717 AGCTATGGCTTCGAGAGAGCATTGCTCTTACCTGTCCCTGACCAAGTGTGAG 767  
 ...  
 RESULT 7  
 US-09-867-550-951  
 ; Sequence 951, Application US/09867550  
 ; Patent No. US20020082206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad,  
 ; APPLICANT: Conley, Pamela  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
 ; THEREBY  
 ; FILE REFERENCE: 21402-013 (Cura-313)  
 ; CURRENT APPLICATION NUMBER: US/09/867,550  
 ; PRIOR FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: USSN 60/208,427  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 951  
 ; LENGTH: 444  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-867-550-951  
 Alignment Scores:  
 Pred. No.: 6,21e-43 Length: 444  
 Score: 452.50 Matches: 96  
 Percent Similarity: 76.80% Conservative: 0  
 Best Local Similarity: 76.80% Mismatches: 29  
 Query Match: 33.44% Indels: 9  
 DB: 1 Gaps: 1  
 US-09-939-853A-75 (1-261) x US-09-867-550-951 (1-444)  
 QY 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 ...

Db 157 ATCGGAAGTCTGCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTGAGTTCTCTCTC 216  
 ...  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 ...  
 Db 217 CAAGGCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGCCACAGCCGCTGGCCCTG 276  
 ...  
 QY 41 GlySerPheProAlaGlyProAlaGluLeuSerLeuArgLeuGluProLeuThr 60  
 ...  
 Db 277 GGCAAGTTCCCGCAGTGGCCCGCCGAGCTGTCTGAGACTCGGGAGGCCATTGACC 336  
 ...  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 ...  
 Db 337 ATCGTCTCTGAG----- 348  
 ...  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 ...  
 Db 349 -----TGGCTGTATGAGGCGCTGAGC 369  
 ...  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 ...  
 Db 370 AGGAGAAACAGAGAACTGTCTGTGTACCTGGGAACCTCGAGGGGCGCTTCTCTCATC 429  
 ...  
 QY 121 ArgGluSerGlnThr 125  
 ...  
 Db 430 CGGAGAGCCAGACC 444  
 ...  
 RESULT 8  
 US-10-175-523-50  
 ; Sequence 50, Application US/10175523  
 ; Publication No. US20030096264A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brockman, Jeffrey  
 ; APPLICANT: Evans, David  
 ; APPLICANT: Hook, Derek  
 ; APPLICANT: Klimczak, Leszek  
 ; APPLICANT: Laeng, Pascal  
 ; APPLICANT: Palfreyman, Michael  
 ; APPLICANT: Rajan, Prithi  
 ; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)  
 ; FILE REFERENCE: 3235/1J795-US3  
 ; CURRENT APPLICATION NUMBER: US/10/175,523  
 ; CURRENT FILING DATE: 2002-06-18  
 ; PRIOR APPLICATION NUMBER: US 60/299,151  
 ; PRIOR FILING DATE: 2001-06-18  
 ; PRIOR APPLICATION NUMBER: US 60/317,828  
 ; PRIOR FILING DATE: 2001-09-07  
 ; PRIOR APPLICATION NUMBER: US 60/325,150  
 ; PRIOR FILING DATE: 2001-09-25  
 ; PRIOR APPLICATION NUMBER: US 60/333,047  
 ; PRIOR FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: US 60/349,936  
 ; PRIOR FILING DATE: 2002-01-18  
 ; PRIOR APPLICATION NUMBER: US 60/361,834  
 ; PRIOR FILING DATE: 2002-03-04  
 ; NUMBER OF SEQ ID NOS: 197  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 50  
 ; LENGTH: 2298  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-175-523-50  
 Alignment Scores:  
 Pred. No.: 1,81e-32 Length: 2298  
 Score: 370.50 Matches: 80  
 Percent Similarity: 57.71% Conservative: 36  
 Best Local Similarity: 59.60% Mismatches: 76  
 Query Match: 27.38% Indels: 9  
 DB: 14 Gaps: 3  
 US-09-939-853A-75 (1-261) x US-10-175-523-50 (1-2298)  
 QY 6 SerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyPro 25  
 ...

```

Db 409 TCCAAATAACAGCAAGCCAGTCCAGAA---TCTCAGCTTTTACCTGGACAGAGGTTT 465
Qy 26 ValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPheProAla 45
Db 466 CAAACTAAAGATCCAGAGGCAAGAGGAGCAATTTGGTAGCCTTTGACCCCTATGATGCG 525
Qy 46 GlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGluAsp 65
Db 526 ATCCACCGGACGACTTGTCTTTCAAGAAAGGAGAGAGATGAAGTCCCTGGAGGAGCAT 585
Qy 66 GlyAspTrpThrValLeuSerGluValSerGlyArgGluThrAsnIleProSerVal 85
Db 586 GGAGAAATGGTGGAAAGCAAGTCCCTTTTAAACAAAAAAGAGAGGCTTCATCCCAAGCAAC 645
Qy 86 HisValGlyLysVal-----SerHisGlyTrpLeuTrpGluGlyLeuSerArg 101
Db 646 TATGTGCCAAACTCAACACTTAGAACACAGAGAGTGGTTTTTCAAGGATATAACACAGG 705
Qy 102 GluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArg 121
Db 706 AAGACCCAGAAAGCGCAGCTTTTGGCACCAGGAAATAGCGCTGGAGCTTTCCTTATTAGA 765
Qy 122 GluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSer 141
Db 766 GAAAGTGAACATTAAGACACTCAAAATAGAAAGTCTGTCTGTCTGACAGACTTTCACCTGTGCAT 825
Qy 142 TrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTrpIleSer 161
Db 826 GGTGATGTTATTAGACACTCAAAATAGAAAGTCTGGATAATGGGGCTATTACATCTCT 885
Qy 162 ProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAsp 181
Db 886 CCAGATACACTTTTCCCTGTATCAGCGCATGATTAACATTAACCAAGAGGAGCAT 945
Qy 182 AspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGly 201
Db 946 GCGTTGTGCAGAAAGTGGAGAGGCTTGTATT-----AGTCCCAAGCCACAG 993
Qy 202 Lys 202
Db 994 AAG 996

```

RESULT 9

```

US-10-159-563-343
; Sequence 343, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 343
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-343

```

```

Alignment Scores:
Pred. No.: 1,818-32 Length: 2298
Score: 370.50 Matches: 80
Percent Similarity: 57.71% Conservative: 36
Best local Similarity: 39.80% Mismatches: 76
Query Match: 27.38% Indels: 9

```

```

DB: 15 Gaps: 3
US-09-939-853A-75 (1-261) x US-10-159-563-343 (1-2298)
Qy 6 SerArgArgLysSerLeuProSerProSerLeuSerSerSerValGlnGlyGlnGlyPro 25
Db 409 TCCAAATAACAGCAAGCCAGTCCAGAA---TCTCAGCTTTTACCTGGACAGAGGTTT 465
Qy 26 ValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPheProAla 45
Db 466 CAAACTAAAGATCCAGAGGCAAGAGGAGCAATTTGGTAGCCTTTGACCCCTATGATGCG 525
Qy 46 GlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGluAsp 65
Db 526 ATCCACCGGACGACTTGTCTTTCAAGAAAGGAGAGAGATGAAGTCCCTGGAGGAGCAT 585
Qy 66 GlyAspTrpThrValLeuSerGluValSerGlyArgGluThrAsnIleProSerVal 85
Db 586 GGAGAAATGGTGGAAAGCAAGTCCCTTTTAAACAAAAAAGAGAGGCTTCATCCCAAGCAAC 645
Qy 86 HisValGlyLysVal-----SerHisGlyTrpLeuTrpGluGlyLeuSerArg 101
Db 646 TATGTGCCAAACTCAACACTTAGAACACAGAGAGTGGTTTTTCAAGGATATAACACAGG 705
Qy 102 GluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArg 121
Db 706 AAGACCCAGAAAGCGCAGCTTTTGGCACCAGGAAATAGCGCTGGAGCTTTCCTTATTAGA 765
Qy 122 GluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSer 141
Db 766 GAAAGTGAACATTAAGACACTCAAAATAGAAAGTCTGTCTGTCTGACAGACTTTCACCTGTGCAT 825
Qy 142 TrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTrpIleSer 161
Db 826 GGTGATGTTATTAGACACTCAAAATAGAAAGTCTGGATAATGGGGCTATTACATCTCT 885
Qy 162 ProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAsp 181
Db 886 CCAGATACACTTTTCCCTGTATCAGCGCATGATTAACATTAACCAAGAGGAGCAT 945
Qy 182 AspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGly 201
Db 946 GCGTTGTGCAGAAAGTGGAGAGGCTTGTATT-----AGTCCCAAGCCACAG 993
Qy 202 Lys 202
Db 994 AAG 996

```

RESULT 10

```

US-10-193-720-1
; Sequence 1, Application US/10193720
; Publication No. US20040009173A1
; GENERAL INFORMATION:
; APPLICANT: Frieria, Annabelle M.
; APPLICANT: Wong, Brian R.
; APPLICANT: Masuda, Etsaban
; APPLICANT: Powell, Mark
; TITLE OF INVENTION: Modulators of Leukocyte Activation, Hck Compositions and Methods
; FILE REFERENCE: A-71313/RMS/DHR
; CURRENT APPLICATION NUMBER: US/10/193,720
; CURRENT FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1595)
; OTHER INFORMATION:
US-10-193-720-1

```

```

Alignment Scores:
Pred. No.: 2,02e-31 Length: 1924
Score: 360.50 Matches: 77
Percent Similarity: 58.38% Conservative: 31
Best Local Similarity: 41.62% Mismatches: 70
Query Match: 26.64% Indels: 7
DB: 15 Gaps: 2

US-09-939-853A-75 (1-261) x US-10-193-720-1 (1-1924)
QY 12 ProSerProSerLeuSerSerValGlnGlyProValThrMetGluAlaGlu 31
DB 195 CCGGGGCTTAATAGCCACACAGC-----AACACACACAGGAATCAGGAGCGGC 245
QY 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeu 51
DB 246 TCTGAGGACATCATCGTGGTTCCTGTATGATGATGAGGCCATTCACACAGAACCTC 305
QY 52 SerLeuArgLeuGlyLysProLeuThrIleValSerGluAspGlyAspTrpThrVal 71
DB 306 AGCTTCACAGAGGGGACACAGATGGTGTCTTAGAGGAATCCGGGGAGTGTGGAAGGCT 365
QY 72 LeuSerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal 90
DB 366 CGATCCCTGGCCACCCGGAAGGAGGCTACATCCCAAGCAACTATGTCCCGCGTTGAC 425
QY 91 -----SerHisGlyTrpLeuTyrGluGlyLeuSerArgGluLysAlaGluGluLeu 107
DB 426 TCTCTGAGACAGAGAGGTGGTTTTCAGGGGCATCAGCCGGAAGGACGACGAGCCCAA 485
QY 108 LeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArg 127
DB 486 CTGCTGGCTCCCGGCAACATGCTGGGCTCTTTCATGATCCGGGATAGGACACCTAAA 545
QY 128 GlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147
DB 546 GGAAGTACTCTTTGTCGCTGGAGACTACGACCTCGCGGAGGAGATACCGTGAACAT 605
QY 148 TyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeuThrPhePro 167
DB 606 TACAAGATCCGGACCTCGACACACGGGGCTTCTACATATCCCCCGAAGACCTTCAGC 665
QY 168 SerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeu 187
DB 666 ATCTTCAGAGCTGTGTGACCACTTACAAGAGGGGGAACGACGGGCTCTGCCAGAACTG 725
QY 188 LysGluProCysVal 192
DB 726 TCGGTGCCCTGCATG 740

RESULT 11
US-09-954-456-1983
; Sequence 1983, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26

```

```

; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1983
; LENGTH: 2015
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-1983

Alignment Scores:
Pred. No.: 2,16e-31 Length: 2015
Score: 360.50 Matches: 77
Percent Similarity: 58.38% Conservative: 31
Best Local Similarity: 41.62% Mismatches: 70
Query Match: 26.64% Indels: 7
DB: 9 Gaps: 2

US-09-939-853A-75 (1-261) x US-09-954-456-1983 (1-2015)
QY 12 ProSerProSerLeuSerSerValGlnGlyProValThrMetGluAlaGlu 31
DB 286 CCGGGGCTTAATAGCCACACAGC-----AACACACACAGGAATCAGGAGCGGC 336
QY 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeu 51
DB 337 TCTGAGGACATCATCGTGGTTCCTGTATGATGATGAGGCCATTCACACAGAACCTC 396
QY 52 SerLeuArgLeuGlyLysProLeuThrIleValSerGluAspGlyAspTrpThrVal 71
DB 397 AGTTCAGAGAGGGGACACAGATGGTGTCTTAGAGGAATCCGGGAGTGTGGAAGGCT 456
QY 72 LeuSerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal 90
DB 457 CGATCCCTGGCCACCCGGAAGGAGGCTACATCCCAAGCAACTATGTCCCGCGTTGAC 516
QY 91 -----SerHisGlyTrpLeuTyrGluGlyLeuSerArgGluLysAlaGluGluLeu 107
DB 517 TCTCTGAGACAGAGAGGTGGTTTTCAGGGGATCAGCCCTCGCGAGGAGATACCGTGAACAT 576
QY 108 LeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArg 127
DB 577 CTGCTGGCTCCCGCAACATGCTGGGCTCTTTCATGATCCGGGATAGCAGACCACTAAA 636
QY 128 GlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147
DB 637 GGAAGTACTCTTTTGTCCGTGCGAGACTACGACCTCGCGAGGAGATACCGTGAACAT 696
QY 148 TyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeuThrPhePro 167
DB 697 TACAAGATCCGGACCTCGACACACGGGGCTTCTACATATCCCCCGAAGACCTTCAGC 756
QY 168 SerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeu 187
DB 757 ACTCTGAGGAGCTGTGTGACCACTTACAAGAGGGGGAACGACGGGCTCTGCCAGAACTG 816
QY 188 LysGluProCysVal 192
DB 817 TCGGTGCCCTGCATG 831

RESULT 12
US-10-007-010-3
; Sequence 3, Application US/10007010
; Publication No. US20030125275A1
; GENERAL INFORMATION:
; APPLICANT: Alexander H. Borchers
; APPLICANT: Kenneth W. Dobie

```

; APPLICANT: Loring, Jeanne F. ; Kaser, Matthew R.  
 ; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
 ; FILE REFERENCE: PA-0026-1 CIP

; CURRENT APPLICATION NUMBER: US/10/062,674  
 ; CURRENT FILING DATE: 2002-01-30  
 ; PRIOR APPLICATION NUMBER: US 09/625,102  
 ; PRIOR FILING DATE: 2000-07-24  
 ; NUMBER OF SEQ ID NOS: 2217  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 2038  
 ; LENGTH: 2343  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20040005559A1 411296.5  
 US-10-062-674-2038

Alignment Scores:  
 Pred. No.: 2,66e-31 Length: 2343  
 Score: 360.50 Matches: 77  
 Percent Similarity: 58.38% Conservative: 31  
 Best Local Similarity: 41.62% Mismatches: 70  
 Query Match: 26.64% Indels: 7  
 DB: 15 Gaps: 2

US-09-939-853A-75 (1-261) x US-10-062-674-2038 (1-2343)

QY 12 ProSerProSerLeuSerSerValGlnGlnGlyProValThrMetGluAlaGlu 31  
 Db 592 CCGGGGCTAATAGCCACACAGC-----AACACACACAGGATCAGGGAGCGCAGC 642  
 QY 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeu 51  
 Db 643 TCTGAGGACATCATCGTGGTTCCTGTATGATACGAGCCATTCACACAGGACCTC 702  
 QY 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrVal 71  
 Db 703 AGCTTCAGAGGGGACAGGATGCTCTAGAGGATCCGGGAGTGGTGGAGGCT 762  
 QY 72 LeuSerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal--- 90  
 Db 763 CGATCCCTGCCACCGGAGGAGGCTACATCCCAAGCAACTATGTCGCCGCGTTCAC 822  
 QY 91 -----SerHisGlyTyrLeuTyrGluGlyLeuSerArgGlyLysAlaGluGluLeu 107  
 Db 823 TCTCTGAGACAGAGGATGTTTTCAGGGGATCAGCCGAGGACGACGAGCCCAA 882  
 QY 108 LeuLeuLeuProGlyAsnProGlyAlaPheLeuLeuArgGluSerGlnThrArgArg 127  
 Db 883 CTGCTGGCTCCCGCAACATGCTGGGCTCTTCATGATCCGGATAGCGAGACCACTAAA 942  
 QY 128 GlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147  
 Db 943 GGAAGCTACTCTCTGCTCCGTCGAGACTACGACCTTCGGCAGGAGATACCGTGAACAT 1002  
 QY 148 TyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIleSerProArgLeuThrPhePro 167  
 Db 1003 TACAAGATCCGACCTGGACCAACGGGGCTTCTACATATCCCGGAGACCTTCAGC 1062  
 QY 168 SerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeu 187  
 Db 1063 ACTCTCAGAGAGTGGTGGACCACTACAAGAGGGGAGCAGCGGGCTCTGCCAGAACTG 1122  
 QY 188 LysGluProCysVal 192  
 Db 1123 TCGGTGCCCTGCATG 1137

RESULT 15

; US-09-917-800A-1611  
 ; Sequence 1611, Application US/09917800A  
 ; Patent No. US20020119462A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mendrick, Donna  
 ; APPLICANT: Porter, Mark  
 ; APPLICANT: Johnson, Kory

; APPLICANT: Castle, Arthur  
 ; APPLICANT: Elashoff, Michael  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Molecular Toxicology Modeling  
 ; FILE REFERENCE: 49321-5038-US  
 ; CURRENT APPLICATION NUMBER: US/09/917,800A  
 ; CURRENT FILING DATE: 2001-07-31  
 ; PRIOR APPLICATION NUMBER: US 60/222,040  
 ; PRIOR FILING DATE: 2000-07-31  
 ; PRIOR APPLICATION NUMBER: US 60/222,880  
 ; PRIOR FILING DATE: 2000-11-02  
 ; PRIOR APPLICATION NUMBER: US 60/290,029  
 ; PRIOR FILING DATE: 2001-05-11  
 ; PRIOR APPLICATION NUMBER: US 60/290,645  
 ; PRIOR FILING DATE: 2001-05-15  
 ; PRIOR APPLICATION NUMBER: US 60/292,336  
 ; PRIOR FILING DATE: 2001-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/295,798  
 ; PRIOR FILING DATE: 2001-06-06  
 ; PRIOR APPLICATION NUMBER: US 60/297,457  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,884  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/303,459  
 ; PRIOR FILING DATE: 2001-07-09  
 ; NUMBER OF SEQ ID NOS: 1740  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1611  
 ; LENGTH: 1911  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_013185  
 US-09-917-800A-1611

Alignment Scores:  
 Pred. No.: 3,41e-31 Length: 1911  
 Score: 358.50 Matches: 93  
 Percent Similarity: 52.22% Conservative: 48  
 Best Local Similarity: 34.44% Mismatches: 94  
 Query Match: 26.50% Indels: 35  
 DB: 9 Gaps: 6

US-09-939-853A-75 (1-261) x US-09-917-800A-1611 (1-1911)

QY 1 MetGlySerLeuProSerArg-----ArgLysSerLeuProSerProSerLeuSerSer 18  
 Db 184 ATGGGATGTGTGAAGTCCAGGTTCCTCGAGAGGAGCAAGGCTCAAAAATAGAGCCA 243  
 QY 19 SerValGlnGlyGlnGlyProValThrMetGlu----- 29  
 Db 244 AATGCCAACCAAGAAAGGCCCTGTGTATGTCCGGATCCACGTCCTCCCTAAGAGCTGGGA 303  
 QY 30 -----AlaGluArgSerLysAlaThr---Ala 37  
 Db 304 CCGAACAGCATCAACACCTGCCCCCGGGTTCGTGGAGGGCTCTGAGGACCACTTGTG 363  
 QY 38 ValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGlyGlu 57  
 Db 364 GTCGCACTGTACGACTATGAGGCGCATTCACCGTGAAGACCTCAGCTTCCAGAGGGAGAC 423  
 QY 58 ProLeuThrIleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGly 77  
 Db 424 CAGATGTGGTTCCTGGAGGAGTCTGGGGAGTGTGGAGGCCGCTTCCCTGGCTTACCAG 483  
 QY 78 ArgGluTyrAsnIleProSerValHisValGlyLysVal-----SerHisGly 93  
 Db 484 AAAGAGGCTATATCCCAAGCAATATGTAGTCCGAGTTAACTCTTTGGAGACTGAGGAG 543  
 QY 94 TrpLeuTyrGluGlyLeuSerArgGluLysAlaGluLeuLeuLeuProGlyAsn 113  
 Db 544 TGGTTCCTCAAGGGTATCAGCCGGAAGGATGTCAGAGCGCCACCTGTGTGCTCCCGGAAC 603

```
QY 114 ProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTySerLeuSer 133
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
604 ATGCTGGGCTCCTTCATGATCCGGACAGTGAGACCAACCAAGGAGGAGCTACTCATTCT 663
QY 134 ValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyArgIleHisCysLeu 153
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
664 GTTCGAGACTTTGACCCGCCAGCAGGACACGGTGAAGCATTATAAATCCGGACACTG 723
QY 154 AspAsnGlyTrpLeuTyrlleSerProArgLeuThrPheProSerLeuGlnAlaLeuVal 173
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
724 GACAGTGGAGGGTTCTACATCTCTCCGAGGAGCACCTTCAGCAGCCTGCAGGAACCTGTC 783
QY 174 AspHisTySerGluLeuAlaAspIleCysCysLeuLeuLysGluProCysVal-Ie 193
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
784 GTCCACTACAAGAAGGAGGAGATGGGCTCTGCCAGAGCTGTCAGTGCCTGTGTCT 843
QY 193 uGln-----ArgAlaGlyProLeuProGlyLysAspileProLeuProValTh 209
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
844 CCGAAACCCAGAACCCATGGGAGAAAGATGCTTGGGAGATTCTCGAGAATCCCTGCAG 903
QY 209 rValGlnArgThr-----ProLeuAsn-TripLysGluLeuAspSerSerLeuLeuPheS 227
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
904 ATGGAGAGAAACTGGGAGCCGGCAGTTTGGAGAGTGTGGATGGCCACCTACAACAAG 963
QY 227 erGluAlaAlaThrGlyGluGluSer 235
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
964 CACACCAAAGTGCGGTGAAGACAAT 989
```

Search completed: February 20, 2004, 12:38:07  
Job time : 370 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:18 ; Search time 87.562 Seconds  
(without alignments)  
9899.970 Million cell updates/sec

Title: US-09-939-853A-140

Perfect score: 20

Sequence: 1 ctggacaggttaggcttgg 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pi.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pi.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rnd.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	6	AX443199	AX443199 Sequence
2	20	100.0	1183	6	AX443133	AX443133 Sequence
3	20	100.0	1183	6	AX443135	AX443135 Sequence
4	20	100.0	2538	9	BC042041	BC042041 Homo sapi
5	20	100.0	2567	6	AX452880	AX452880 Sequence
6	20	100.0	2788	6	AX780857	AX780857 Sequence
7	20	100.0	65741	9	HS460038	AL031862 Human DNA
8	20	100.0	145833	2	AC026539	AC026539 Homo sapi
9	17.4	87.0	80828	9	AC093247	AC093247 Homo sapi
10	17.4	87.0	100300	2	AC114970	AC114970 Homo sapi
11	17.4	87.0	108742	2	AC023817	AC023817 Homo sapi
12	17.4	87.0	110000	2	AC139485_1	Continuation (2 of
13	17.4	87.0	115702	9	AC132801	AC132801 Homo sapi
14	17.4	87.0	129218	2	AC138851	AC138851 Homo sapi
15	17.4	87.0	133181	2	AC138930	AC138930 Homo sapi
16	17.4	87.0	140410	2	AC139483	AC139483 Homo sapi
17	17.4	87.0	145493	2	AC138839	AC138839 Homo sapi
18	17.4	87.0	146061	9	AC118459	AC118459 Homo sapi
19	17.4	87.0	146175	2	AL161433	AL161433 Homo sapi
20	17.4	87.0	147840	2	AC138939	AC138939 Homo sapi
21	17.4	87.0	149397	9	AC138827	AC138827 Homo sapi
22	17.4	87.0	152445	9	AC138410	AC138410 Homo sapi
23	17.4	87.0	152782	9	AL133255	AL133255 Human DNA
24	17.4	87.0	154803	9	AC131392	AC131392 Homo sapi
25	17.4	87.0	155537	2	AC139282	AC139282 Homo sapi
26	17.4	87.0	156527	2	AC138968	AC138968 Homo sapi
27	17.4	87.0	158755	2	AC145032	AC145032 Homo sapi
28	17.4	87.0	159184	2	AC139484	AC139484 Homo sapi
29	17.4	87.0	160994	2	AC044800	AC044800 Homo sapi
30	17.4	87.0	161000	2	AC145100	AC145100 Homo sapi
31	17.4	87.0	161279	2	AC139508	AC139508 Homo sapi
32	17.4	87.0	161460	9	AC093243	AC093243 Homo sapi
33	17.4	87.0	161460	9	AL157879	AL157879 Human DNA
34	17.4	87.0	161671	9	AC140175	AC140175 Homo sapi
35	17.4	87.0	161893	9	AC137781	AC137781 Homo sapi
36	17.4	87.0	161957	2	AC138909	AC138909 Homo sapi
37	17.4	87.0	162703	9	AC132802	AC132802 Homo sapi
38	17.4	87.0	165211	2	AC145128	AC145128 Homo sapi
39	17.4	87.0	166615	9	AC138928	AC138928 Homo sapi
40	17.4	87.0	167663	2	AC145135	AC145135 Homo sapi
41	17.4	87.0	168547	2	AC138947	AC138947 Homo sapi
42	17.4	87.0	168814	9	AC010237	AC010237 Homo sapi
43	17.4	87.0	169082	2	AC138933	AC138933 Homo sapi
44	17.4	87.0	170336	2	AC011244	AC011244 Homo sapi
45	17.4	87.0	171363	9	AC145129	AC145129 Homo sapi

# ALIGNMENTS

RESULT 1  
AX443199  
LOCUS AX443199  
DEFINITION Sequence 140 from Patent WO0216599.  
ACCESSION AX443199  
VERSION AX443199.1 GI:21690594  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1  
AUTHORS Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R.,  
Shinkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E.,  
Topper, J.N. and Yang, R.B.  
TITLE Proteins and nucleic acids encoding same

Pred. No. is the number of results predicted by chance to have a



JOURNAL Patent: WO 0216599-A 140 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)

## FEATURES

source  
1. .20  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="oligonucleotide primer"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20  
|||||  
Db 1 CTGGACAGGTTAGGCGTTTG 20

## RESULT 2

AX443133/c  
LOCUS AX443133 1183 bp DNA linear PAT 02-JUL-2002  
DEFINITION Sequence 74 from Patent WO0216599.  
ACCESSION AX443133  
VERSION AX443133.1 GI:21690555

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R.,

Shinkets, R.A., Spyttek, K.A., Szekeres, E.S., Tomlinson, J.E.,

Topper, J.N. and Yang, R.B.

TITLE Proteins and nucleic acids encoding same

JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;

Curagen Corporation (US); COR THERAPEUTICS, INC. (US)

FEATURES Location/Qualifiers

1. .1183

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20  
|||||  
Db 301 CTGGACAGGTTAGGCGTTTG 282

## RESULT 3

AX443135  
LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002  
DEFINITION Sequence 76 from Patent WO0216599.  
ACCESSION AX443135  
VERSION AX443135.1 GI:21690556

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R.,

Shinkets, R.A., Spyttek, K.A., Szekeres, E.S., Tomlinson, J.E.,

Topper, J.N. and Yang, R.B.

TITLE Proteins and nucleic acids encoding same

JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;

Curagen Corporation (US); COR THERAPEUTICS, INC. (US)

FEATURES Location/Qualifiers

1. .1183

/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20  
|||||  
Db 883 CTGGACAGGTTAGGCGTTTG 902

## RESULT 4

BC042041/c

LOCUS BC042041 2538 bp mRNA linear PRI 07-OCT-2003

DEFINITION Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (cdna

clone MGC:49845 IMAGE:4429896), complete cds.

ACCESSION BC042041

VERSION BC042041.1 GI:27469842

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mialahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Rulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Boutard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,

Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

2 (bases 1 to 2538)

Strausberg, R.

Direct Submission

Submitted (23-DEC-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapps-remail.nih.gov](mailto:cgapps-remail.nih.gov)

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 88 Row: a Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28416422.

```

FEATURES
  source
    1. .2538
      /location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="MGC:49845 IMAGE:4429896"
        /tissue_type="Prostate, adenocarcinoma."
        /clone_lib="NIH_MGC_91"
        /lab_host="DH10B"
        /note="Vector: pCMV-SPORT6"
      1. .2538
        /gene="SLA2"
        /note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845,
        SLAP2"
        /db_xref="LocusID:84174"
        /db_xref="MIM:606577"
        363. .1148
          /codon_start=1
          /product="Src-like-adaptor 2, isoform a"
          /protein_id="AAH42041.1"
          /db_xref="GI:27469843"
          /db_xref="LocusID:84174"
          /translation="MGSLEPRKSLPSPSSVQGGPVVMEASRKATAVALGSPF
          AGPAELSLRLGEPLTVSDGWTVLSEVGRYNPVSHVAKVSHGWLVEGLSRE
          KAEELLPLGNPGGAPLIRSGYSLSVLSFASWDRIHRHRIHGLDNGWLYI
          SPLTTPSLOALVDHYSELADICCLIKPCVLQAGPLPGKIDPLPVTVQRTPLNWK
          ELSSSLPSEANATGEESLLSEGLRESLSFSLNDEAVSLDDA"
        468. .632
          /note="SH3; Region: SH3 domain. SH3 (Src homology 3)
          domains are often indicative of a protein involved in
          signal transduction related to cytoskeletal organization.
          First described in the Src cytoplasmic tyrosine kinase.
          The structure is a partly opened beta barrel"
          /db_xref="CDD:pfam00018"
        642. .890
          /note="SH2; Region: SH2 domain"
          /db_xref="CDD:pfam00017"

  CDS
    1. .2538
      /product="Src-like-adaptor 2, isoform a"
      /protein_id="AAH42041.1"
      /db_xref="GI:27469843"
      /db_xref="LocusID:84174"
      /translation="MGSLEPRKSLPSPSSVQGGPVVMEASRKATAVALGSPF
      AGPAELSLRLGEPLTVSDGWTVLSEVGRYNPVSHVAKVSHGWLVEGLSRE
      KAEELLPLGNPGGAPLIRSGYSLSVLSFASWDRIHRHRIHGLDNGWLYI
      SPLTTPSLOALVDHYSELADICCLIKPCVLQAGPLPGKIDPLPVTVQRTPLNWK
      ELSSSLPSEANATGEESLLSEGLRESLSFSLNDEAVSLDDA"
      468. .632
        /note="SH3; Region: SH3 domain. SH3 (Src homology 3)
        domains are often indicative of a protein involved in
        signal transduction related to cytoskeletal organization.
        First described in the Src cytoplasmic tyrosine kinase.
        The structure is a partly opened beta barrel"
        /db_xref="CDD:pfam00018"
      642. .890
        /note="SH2; Region: SH2 domain"
        /db_xref="CDD:pfam00017"

  misc_feature
    1. .2538
      /note="SH3; Region: SH3 domain. SH3 (Src homology 3)
      domains are often indicative of a protein involved in
      signal transduction related to cytoskeletal organization.
      First described in the Src cytoplasmic tyrosine kinase.
      The structure is a partly opened beta barrel"
      /db_xref="CDD:pfam00018"
    642. .890
      /note="SH2; Region: SH2 domain"
      /db_xref="CDD:pfam00017"

  ORIGIN
    Query Match 100.0%; Score 20; DB 9; Length 2538;
    Best Local Similarity 100.0%; Pred. No. 4.4;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1 CTGCACAGGTTAGGCGTTTG 20
    DB 266 CTGCACAGGTTAGGCGTTTG 247

  RESULT 5
  LOCUS AX452880 2567 bp DNA linear PAT 06-JUL-2002
  DEFINITION Sequence 1 from Patent WO0242457.
  ACCESSION AX452880
  VERSION AX452880.1 GI:21712520
  KEYWORDS Homo sapiens (human)
  SOURCE Homo sapiens
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villair, J.J. and
    Kanner, S.B.
  AUTHORS Cloning and expression of human slap-2: a novel sh2/sh3
    domain-containing human slap homologue having immune cell-specific
    expression
  TITLE Patent: WO 0242457-A 1 30-MAY-2002;
  JOURNAL Bristol-Myers Squibb Co. (US)
  FEATURES
    source
      1. .2567
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"

```

```

  ORIGIN
    /db_xref="taxon:9606"

    Query Match 100.0%; Score 20; DB 6; Length 2567;
    Best Local Similarity 100.0%; Pred. No. 4.4;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1 CTGCACAGGTTAGGCGTTTG 20
    DB 318 CTGCACAGGTTAGGCGTTTG 299

  RESULT 6
  LOCUS AX780857/c 2788 bp DNA linear PAT 14-JUL-2003
  DEFINITION Sequence 3014 from Patent WO03039443.
  ACCESSION AX780857
  VERSION AX780857.1 GI:32697851
  KEYWORDS Homo sapiens (human)
  SOURCE Homo sapiens
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE Haeflrich, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
    Dugas, M., Ellis, R., Brots, B. and Mergenthaler, S.
  AUTHORS Novel genetic markers for leukemias
  TITLE Patent: WO 03039443-A 3014 15-MAY-2003;
  JOURNAL Deutsches Krebsforschungszentrum (DE);
    Ludwig-Maximilian-Universitaet Muenchen (DE); Haeflrich, Torsten,
    PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)

  FEATURES
    source
      1. .2788
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

  ORIGIN
    Query Match 100.0%; Score 20; DB 6; Length 2788;
    Best Local Similarity 100.0%; Pred. No. 4.4;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1 CTGCACAGGTTAGGCGTTTG 20
    DB 290 CTGCACAGGTTAGGCGTTTG 271

  RESULT 7
  LOCUS HS460J8 66741 bp DNA linear PRI 23-JUL-2001
  DEFINITION Human DNA sequence from clone RP3-460J8 on chromosome
    20q11.21-11.23 Contains the 3' end of the gene for a novel protein
    similar to N-myc downstream regulated (NDRG1) the 5' end of a gene
    encoding a novel protein tyrosine kinase, ESTs, STSS and GSSs,
    complete sequence.
  ACCESSION AL031662
  VERSION AL031662.26 GI:9716901
  KEYWORDS HTG; NDRG1; SH2 domain.
  SOURCE Homo sapiens (human)
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 66741)
  AUTHORS Skuce, C.
  TITLE Direct Submission
  JOURNAL Submitted (23-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
    CB10 1BA, UK. E-mail enquiries: humquery@sanger.ac.uk
    requests: clonerequest@sanger.ac.uk
  COMMENT On Aug 7, 2000 this sequence version replaced gi:6425549.
    During sequence assembly data is compared from overlapping clones.
    Where differences are found these are annotated as variations
    together with a note of the overlapping clone name. Note that the
    variation annotation may not be found in the sequence submission

```

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human

chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr20>

IMPORTANT: This sequence is not the entire insert of clone

RP3-460J8 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP3-460J8 is at 66741 in this sequence. The true left end of clone RP3-460J8 is at 41767 in this sequence.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. RP3-460J8 is from the library RPCI-3 constructed by the group of Pieter de Jong. For

further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

#### FEATURES

```

source
    location/Qualifiers
    1..66741
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="20"
    /map="GII-21-11-23"
    /clone="RP3-460J8"
    /clone_lib="RPCI-3"
    /complement(50..544)
    /note="match: GSS: Em:B45150"
    /complement(240..7128)
    /gene="dJ977B1.1"
    /complement(join(<240..339,6995..7128))
    /gene="dJ977B1.1"
    /product="dJ460J8.2 (novel protein tyrosine kinase with
    Src homology 2 (SH2) domain)"
    /note="match: cDNAs: Em:AK025645
    match: ESTs: Em:BG178487"
    /evidence=not experimental
    /complement(join(<240..339,6995..7085))
    /gene="dJ977B1.1"
    /note="Continues in Em:AL050318 as dJ977B1.2"
    /codon_start=1
    /evidence=not experimental
    /product="dJ460J8.2 (novel protein tyrosine kinase with
    Src homology 2 (SH2) domain)"
    /protein_id="CAC44645.1"
    /db_xref="GI:15020830"
    /db_xref="GOA:Q9H6Q3"
    /db_xref="SWISS-PROT:Q9H6Q3"
    /translation="MGSLPSRRKSLPSLSSVQGGQPVVTMEATERSKATAVALGSPFP
    AGGPAELSLRGLPELTIVS"
    536..565
    /note="15 copies 2 mer ag 93% conserved"
    repeat_region
    2375..2402
    /note="14 copies 2 mer ta 100% conserved"
    2783..3155
    /note="match: GSS: Em:AQ807191"
    /complement(6983..7482)
    /note="match: GSS: Em:AQ556467"
    /complement(5985..7492)
    /note="match: GSS: Em:AQ556478"
    7486..7925
    /note="match: STS: Em:HS427J1S"

```

```

repeat_region
    10615..10640
    /note="13 copies 2 mer tt 92% conserved"
    13593..13961
    /note="MER21B repeat: matches 422..785 of consensus"
    14248..14297
    /note="MER21B repeat: matches 374..422 of consensus"
    14594..14614
    /note="MER21B repeat: matches 355..374 of consensus"
    14746..15078
    /note="MER21B repeat: matches 1..355 of consensus"
    /complement(17516)
    17519..17686
    /note="match: GSS: Em:AZ067993"
    /complement(17521)
    /complement(17522)
    /complement(17523..54534)
    /gene="dJ469A13.3"
    /complement(join(<17523..19451,20551..20589,22110..22161,
    22247..22282,26087..26134,30796..30847,30986..31051,
    32039..32142,37091..37147,46566..46652,48268..48328,
    50163..50225,53242..53362,54429..54534))
    /gene="dJ469A13.3"
    /product="dJ460J8.1 (continued from dJ469A13.3 in
    Em:AL132768)"
    /note="match: cDNAs: Em:AB033922 Em:U52073 Em:D87953
    Em:X92845 Em:AF004162 Em:U60593 Em:AF045564 Em:AB033921
    Em:AF159092 Em:AF147402 Em:M59814
    match: ESTs: Em:AA718726 Em:AA039000 Em:AW003952 Em:W89263
    Em:AA325826 Em:AI230982 Em:AA162360 Em:AA445016
    Em:AV002395 Em:T85147 Em:AI786615 Em:AI13437 Em:AI004026
    Em:T88705 Em:AI786673 Em:AI181197 Em:AA764653 Em:AA316771
    Em:AU035165 Em:AV002368"
    /evidence=not experimental
    /complement(17542..17547)
    /gene="dJ469A13.3"
    17555..17811
    /note="match: STS: Em:G19945"
    /join(19202..19451,22110..22282,26087..26127,30796..30847)
    /note="match: STS: Em:G23762"
    /complement(join(19270..19451,20551..20589,22110..22161,
    22247..22282,26087..26134,30796..30847,30986..31051,
    32039..32142,37091..37147,46566..46652,48268..48328,
    50163..50225,53242..53362,54429..54534))
    /gene="dJ469A13.3"
    /note="novel protein (FLJ13556) similar to N-myc
    downstream regulated (NDRG1)
    match: proteins: Sw:Q62433 Sw:Q92597 Tr:Q922L9 Sw:P97862"
    /codon_start=1
    /evidence=not experimental
    /product="dJ460J8.1 (continued from dJ469A13.3 in
    Em:AL132768)"
    /protein_id="CAB65625.1"
    /db_xref="GI:6687781"
    /db_xref="GOA:Q9UGV2"
    /db_xref="SWISS-PROT:Q9UGV2"
    /translation="EHDLETHGVVHTIRGLPKGNRPVILTYHDIGLNHKSCEVNAFF
    NFEDMQEITQHFVAVCHVDPAPGQEGAPFPPTGYQVPTDELAEMLPVLTSLKSLII
    GIGVGAGYILSRFALNHPGLVGLVINDPCAKGNDWAASKLSGLTINVDLILA
    HFHFGEEQLANDLIQTYRMHIAQDINQDLFLNSYNGRDLIERPILQGNOKS
    KTLKCSLTLLVGDNSPAVEAVVECNRLNPITLLKRWADCGLPQVPGKLTAFK
    YFLQMGVTPYVQLSHLSTESVPSASMTLARSRTHTSTSSLSGSGESFPFSRVTSNQS
    DGTQSCESPVDLDRHQTMVEVC"
    /complement(24246..24780)
    /gene="dJ469A13.3"
    /note="match: GSS: Em:AQ592789"
    /complement(27980..28556)
    /gene="dJ469A13.3"
    /note="match: GSS: Em:AQ308867"
    34313..34420
    /note="HY1 repeat: matches 1..109 of consensus"
    34505..34560
    /note="28 copies 2 mer ta 78% conserved"
    34564..34611

```

```

repeat_region /note="24 copies 2 mer ta 79% conserved"
42553..43658
/note="MER52C repeat: matches 1. 1278 of consensus"
47989..49470
/note="match: GSS: Em:AQ672417"
49391..49480
/note="45 copies 2 mer ta 76% conserved"
52922..53120
/note="match: STS: Em:G04621"
/note="match: STS: Em:G04621"
complement(53426..54032)
/note="match: GSS: Em:AQ314824
match: STS: Em:G55853"
54054..54477
/note="match: GSS: Em:AQ310681"
57509..57838
/note="match: STS: Em:G07504"
complement(57696..58013)
/note="match: STS: Em:G07632 Em:G07634"
58856..59001
/note="73 copies 2 mer at 83% conserved"
complement(58930..59281)
/note="match: GSS: Em:AQ067563"
59062..59236
/note="TIGER2 repeat: matches 2541. 2714 of consensus"
59738..60495
/note="TIGER2 repeat: matches 1780. 2541 of consensus"
60719..61831
/note="MER11C repeat: matches 1. 1071 of consensus"
62547..63174

```

```

Query Match 100.0%; Score 20; DB 9; Length 66741;
Best Local Similarity 100.0%; Pred. NO. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGTAGGGCTTGG 20
Db 11640 CTGCACAGGTAGGGCTTGG 11659
|||||

```

```

RESULT 8
AC026539/c 145833 bp DNA linear HTG 27-APR-2000
LOCUS Homo sapiens chromosome 20 clone RP11-712N14 map 20, WORKING DRAFT
DEFINITION SEQUENCE, 37 unordered pieces.
ACCESSION AC026539
VERSION AC026539.2 GI:7656813
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 145833)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,I., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castelle,A., Choepe,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lenoczky,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McKernan,K., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mhova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

```

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 27, 2000 this sequence version replaced gi:7283243.  
 All repeats were identified using RepeatMasker:  
 Smit,A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center -----

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIER  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information -----  
 Center project name: L7115  
 Center clone name: 712.N.14  
 ----- Summary Statistics -----

Sequencing vector: M13; M7815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 125577 bases at least Q40  
 Consensus quality: 135703 bases at least Q30  
 Consensus quality: 139593 bases at least Q20  
 Insert size: 182000; agarose-fp  
 Insert size: 142233; sum-of-contigs  
 Quality coverage: 2.6 in Q20 bases; agarose-fp  
 Quality coverage: 3.3 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 37 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

* 1 1214: contig of 1214 bp in length
* 1215 1314: gap of 100 bp
* 1315 2673: contig of 1359 bp in length
* 2674 2773: gap of 100 bp
* 2774 4520: contig of 1747 bp in length
* 4521 4620: gap of 100 bp
* 4621 5961: contig of 1341 bp in length
* 5962 6061: gap of 100 bp
* 6062 7719: contig of 1658 bp in length
* 7720 7819: gap of 100 bp
* 7820 9799: contig of 1980 bp in length
* 9800 9899: gap of 100 bp
* 9900 11434: contig of 1535 bp in length
* 11435 11534: gap of 100 bp
* 11535 14382: contig of 2848 bp in length
* 14383 14482: gap of 100 bp
* 14483 17148: contig of 2666 bp in length
* 17149 17248: gap of 100 bp
* 17249 19239: contig of 1991 bp in length
* 19240 19339: gap of 100 bp
* 19340 21102: contig of 1763 bp in length
* 21103 21202: gap of 100 bp
* 21203 23371: contig of 2169 bp in length
* 23372 23471: gap of 100 bp
* 23472 25782: contig of 2311 bp in length
* 25783 25882: gap of 100 bp
* 25883 28824: contig of 2942 bp in length
* 28825 28924: gap of 100 bp
* 28925 31619: contig of 2695 bp in length
* 31620 31719: gap of 100 bp
* 31720 34679: contig of 2960 bp in length
* 34680 34779: gap of 100 bp
* 34780 37527: contig of 2748 bp in length

```

\* 37528 37627: gap of 100 bp  
\* 37628 40146: contig of 2519 bp in length  
\* 40147 40246: gap of 100 bp  
\* 40247 43743: contig of 3497 bp in length  
\* 43744 43843: gap of 100 bp  
\* 43844 47133: contig of 3290 bp in length  
\* 47134 47233: gap of 100 bp  
\* 47234 51023: contig of 3790 bp in length  
\* 51024 51123: gap of 100 bp  
\* 51124 54935: contig of 3812 bp in length  
\* 54936 55035: gap of 100 bp  
\* 55036 59553: contig of 4518 bp in length  
\* 59554 59653: gap of 100 bp  
\* 59654 64428: contig of 4775 bp in length  
\* 64429 64528: gap of 100 bp  
\* 64529 69211: contig of 4583 bp in length  
\* 69212 69311: gap of 100 bp  
\* 69312 72901: contig of 3590 bp in length  
\* 72902 73001: gap of 100 bp  
\* 73002 76724: contig of 3723 bp in length  
\* 76725 76824: gap of 100 bp  
\* 76825 81179: contig of 4355 bp in length  
\* 81180 81279: gap of 100 bp  
\* 81280 87009: contig of 5730 bp in length  
\* 87010 87109: gap of 100 bp  
\* 87110 90855: contig of 3746 bp in length  
\* 90856 90955: gap of 100 bp  
\* 90956 96520: contig of 5565 bp in length  
\* 96521 96620: gap of 100 bp  
\* 96621 102321: contig of 5701 bp in length  
\* 102322 102421: gap of 100 bp  
\* 102422 108293: contig of 5872 bp in length  
\* 108294 108393: gap of 100 bp  
\* 108394 116689: contig of 8296 bp in length  
\* 116690 116789: gap of 100 bp  
\* 116790 125264: contig of 8475 bp in length  
\* 125265 125364: gap of 100 bp  
\* 125365 136354: contig of 10990 bp in length  
\* 136355 136454: gap of 100 bp  
\* 136455 145833: contig of 9379 bp in length.

FEATURES  
source

1. 145833  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20"  
/clone="RP11-712N14"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. 1214  
/note="assembly\_fragment"  
1315..2673  
/note="assembly\_fragment"  
2774..4520  
/note="assembly\_fragment"  
4621..5961  
/note="assembly\_fragment"  
6062..7719  
/note="assembly\_fragment"  
7820..9799  
/note="assembly\_fragment"  
9900..11434  
/note="assembly\_fragment"  
11535..14382  
/note="assembly\_fragment"  
14483..17148  
/note="assembly\_fragment"  
17249..19239  
/note="assembly\_fragment"  
19340..21102  
/note="assembly\_fragment"  
21203..23371  
/note="assembly\_fragment"

misc\_feature 23472..25782  
/note="assembly\_fragment"  
misc\_feature 25883..28824  
/note="assembly\_fragment"  
misc\_feature 28925..31619  
/note="assembly\_fragment"  
misc\_feature 31720..34679  
/note="assembly\_fragment"  
misc\_feature 34780..37527  
/note="assembly\_fragment"  
misc\_feature 37628..40146  
/note="assembly\_fragment"  
misc\_feature 40247..43743  
/note="assembly\_fragment"  
misc\_feature 43844..47133  
/note="assembly\_fragment"  
misc\_feature 47234..51023  
/note="assembly\_fragment"  
misc\_feature 51124..54935  
/note="assembly\_fragment"  
misc\_feature 55036..59553  
vector\_side:right

Query Match 100.0%; Score 20; DB 2; Length 145833;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTAGGCGTTTG 20  
|||||  
Db 100045 CTGGACAGGTAGGCGTTTG 100026

## RESULT 9

AC093247 AC093247 80828 bp DNA linear PRI 26-FEB-2002  
LOCUS Homo sapiens chromosome 5 clone RP11-138M1, complete sequence.  
DEFINITION AC093247  
ACCESSION AC093247.3 GI:18921295  
VERSION AC093247.3  
KEYWORDS HTG.

SOURCE  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 80828)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Unpublished  
DOE Joint Genome Institute.  
2 (bases 1 to 80828)  
Direct Submission  
DOE Joint Genome Institute.  
3 (bases 1 to 80828)  
Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Feb 26, 2002 this sequence version replaced gi:15290461.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.5% of Sequence;  
Estimated Total Number of Errors is 0.4.

## REFERENCE

AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
On Feb 26, 2002 this sequence version replaced gi:15290461.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.5% of Sequence;  
Estimated Total Number of Errors is 0.4.  
NOTE: This insert is not the entire sequence of the clone (entire  
sequence is 147.2kb). It is clipped at the overlap with AC010237.  
The number of bases overlapped is 38259.

## FEATURES

source  
1. 80828  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-388B4"
/clone_lib="RPCI human BAC library 11"

ORIGIN
Query Match      87.0%; Score 17.4; DB 9; Length 80828;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTGGACAGGTTAGGCTTT 19
      |||||
DB      71309 CTGGACAGTCAGGCTTT 71327

RESULT 10
AC114970      100300 bp      DNA      linear      HTG 14-MAR-2002
LOCUS      Homo sapiens chromosome 5 clone RP11-388B4, WORKING DRAFT SEQUENCE,
DEFINITION      7 unordered pieces.
AC114970
AC114970.1 GI:19424431
VERSION      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      DOE Joint Genome Institute.
TITLE      Sequencing of Human Chromosome 5
JOURNAL      Unpublished
REFERENCE
AUTHORS      DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL      Direct Submission
COMMENT
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 555385
Center clone name: RPCI-11_388B4
-----
Summary Statistics
Consensus quality: 95359 bases at least Q40
Consensus quality: 97011 bases at least Q30
Consensus quality: 97825 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Quality coverage: 9.62 in Q20 bases; agarose-fp estimation
Quality coverage: 16.88 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1595: contig of 1595 bp in length
* 1696: gap of unknown length
* 1696: contig of 1499 bp in length
* 3195: gap of unknown length
* 3295: contig of 4404 bp in length
* 7693: gap of unknown length
* 7793: contig of 11003 bp in length
* 18901: gap of unknown length
* 18902: contig of 13618 bp in length
* 32520: gap of unknown length
* 32620: contig of 4976 bp in length
* 37595: gap of unknown length
* 37695: contig of 62605 bp in length.
* 37695: 100300: contig of 62605 bp in length.
Location/Qualifiers

```

```

1. 100300
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-388B4"
/clone_lib="RPCI human BAC library 11"

ORIGIN
Query Match      87.0%; Score 17.4; DB 2; Length 100300;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CTGGACAGGTTAGGCTTT 19
      |||||
DB      64854 CTGGACAGTCAGGCTTT 64872

RESULT 11
AC023817      108742 bp      DNA      linear      HTG 07-MAR-2002
LOCUS      Homo sapiens chromosome 5 clone CTD-2565N12, WORKING DRAFT
DEFINITION      SEQUENCE, 3 ordered pieces.
AC023817
AC023817.5 GI:19224766
VERSION      HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      DOE Joint Genome Institute.
TITLE      Sequencing of Human Chromosome 5
JOURNAL      Unpublished
REFERENCE
AUTHORS      DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL      Direct Submission
COMMENT
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 815265
Center clone name: CTD-El_2565N12
-----
Summary Statistics
Consensus quality: 107872 bases at least Q40
Consensus quality: 108273 bases at least Q30
Consensus quality: 108423 bases at least Q20
Estimated insert size: 137000; agarose-fp estimation
Quality coverage: 8.13 in Q20 bases; agarose-fp estimation
Quality coverage: 10.27 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 31413: contig of 31413 bp in length
* 31414: gap of unknown length

```

\* 31514 101416: contig of 69903 bp in length  
\* 101417 101516: gap of unknown length  
\* 101517 108742: contig of 7226 bp in length.

## FEATURES

Location/Qualifiers  
1..108742  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTD-2565N12"  
/clone\_lib="CalTech human BAC library D"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 108742;  
Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTT 19

|||||  
Db 70123 CTGGACAGTCAGGGCTTT 70141

## RESULT 12

AC139485.1

WPCOMMENT

Sequence split into 4 fragments LOCUS AC139485 Accession AC139485  
Fragment Name Begin End  
AC139485\_0 1 110000  
AC139485\_1 100001 210000  
AC139485\_2 310000 310000  
AC139485\_3 300001 368755

Continuation (2 of 4) of AC139485 from base 100001 (AC139485 Homo sapiens chromosome 5 c  
Query Match 87.0%; Score 17.4; DB 2; Length 110000;  
Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTT 19

|||||  
Db 29275 CTGGACAGTCAGGGCTTT 29293

## RESULT 13

AC132801/c

LOCUS

AC132801 116702 bp DNA linear PRI 13-NOV-2002

DEFINITION Homo sapiens chromosome 5 clone CTC-202015, complete sequence.

AC132801

AC132801

VERSION

AC132801.2 GI:24942889

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 116702)

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission

Unpublished

2 (bases 1 to 116702)

DOE Joint Genome Institute.

Direct Submission

TITLE

Submitted (04-SEP-2002) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

3 (bases 1 to 116702)

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission

Submitted (13-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell

Drive, Walnut Creek, CA 94598, USA

On Nov 13, 2002 this sequence version replaced gi:22711579.

Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center

www.shgc.stanford.edu

Quality: Phrap Quality >=40 100% of Sequence;

Estimated Total Number of Errors is 0.

## FEATURES

Location/Qualifiers  
1..116702  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTC-202015"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 116702;  
Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTT 19

|||||  
Db 60842 CTGGACAGGTCAGGGCTTT 60824

## RESULT 14

AC138851/c

LOCUS

DEFINITION

AC138851

AC138851

VERSION

AC138851.1 GI:27805263

KEYWORDS

HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 129218)

DOE Joint Genome Institute.

Sequencing of Human Chromosome 5

Unpublished

2 (bases 1 to 129218)

DOE Joint Genome Institute.

Direct Submission

Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

-----

Project Information

Center Project Name: 2723051

Center clone name: RPCI-11\_1310P17

-----

Summary Statistics

Consensus quality: 124109 bases at least Q40

Consensus quality: 125106 bases at least Q30

Consensus quality: 125993 bases at least Q20

Estimated insert size: 175000; agarose-fp estimation

Estimated insert size: 128018; sum-of-contigs estimation

Quality coverage: 8.8 in Q20 bases; agarose-fp estimation

Quality coverage: 12.03 in Q20 bases; sum-of-contigs estimation.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 13 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1550: contig of 1550 bp in length

\* 1551: gap of unknown length

\* 1551: contig of 6079 bp in length

\* 7729: gap of unknown length

\* 7829: contig of 1968 bp in length

\* 7830: gap of unknown length

\* 9797: gap of unknown length

\* 9798: contig of 2662 bp in length

\* 9998: gap of unknown length

\* 12559: contig of 2662 bp in length

\* 12560: gap of unknown length

RESULT 15	ACI38930	13181 bp	DNA	linear	HTG 21-JAN-2003
LOCUS	ACI38930	13181 bp	DNA	linear	HTG 21-JAN-2003
DEFINITION	ACI38930	13181 bp	DNA	linear	HTG 21-JAN-2003
ACCESSION	ACI38930	13181 bp	DNA	linear	HTG 21-JAN-2003
VERSION	ACI38930.1	GI:27805342			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	(Bases 1 to 133181)			
AUTHORS	DOE Joint Genome Institute.				
TITLE	Sequencing of Human Chromosome 5				
JOURNAL	Unpublished				
REFERENCE	2	(Bases 1 to 133181)			
AUTHORS	DOE Joint Genome Institute.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint				
COMMENT	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
	-----Genome Center				
	Center: Joint Genome Institute				
	Center Code: JGI				
	Web site: <a href="http://www.jgi.doe.gov">http://www.jgi.doe.gov</a>				
	-----				
	Project Information				
	Center Project Name: 1555206				
	Center clone name: RPCI-11_678014				
	-----				
	Summary Statistics				
	Consensus quality: 131745 bases at least Q40				
	Consensus quality: 131928 bases at least Q30				
	Consensus quality: 132052 bases at least Q20				
	Estimated insert size: 175000; agarose-fp estimation				
	Estimated insert size: 132781; sum-of-contigs estimation				
	Quality coverage: 13.96 in Q20 bases; agarose-fp estimation				



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:15:48 ; Search time 18.5292 Seconds  
(without alignments)  
4585.415 Million cell updates/sec

Title: US-09-939-853A-140

Perfect score: 20

Sequence: 1 ctggacaggttagggcttgg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_29Jan04.\*

1: geneseq1980s.\*

2: geneseq1990s.\*

3: geneseq2000s.\*

4: geneseq2001as.\*

5: geneseq2001bs.\*

6: geneseq2002s.\*

7: geneseq2003as.\*

8: geneseq2003bs.\*

9: geneseq2003cs.\*

10: geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	6	Abk61505 Human NOV
2	20	100.0	444	6	Abq98669 Human ORF
3	20	100.0	445	5	Aas74747 DNA encod
4	20	100.0	763	6	Abq98670 Human ORF
5	20	100.0	1183	6	Abk61465 Human ORF
6	20	100.0	2567	6	Abk43980 Human ORF
7	16.8	84.0	538	6	Add22694 Filamento
8	16.8	84.0	589	7	Abz56309 Aspergill
9	16.8	84.0	809	3	Aaf14531 Aspergill
10	16.8	84.0	2305	7	Adas3227 Human cod
11	16.8	84.0	44325	8	Ada02960 Human DAD
12	16.8	84.0	44325	9	Adb72698 Human DAD
13	16.8	84.0	44325	9	Adc85440 Human Dad
14	16.4	82.0	2446	9	Ades7896 Human gen
15	16.4	82.0	2446	9	Ades7892 Human gen
16	16.4	82.0	2446	9	Ades7900 Human gen
17	16.4	82.0	2446	9	Ades7904 Human gen
18	16.4	82.0	2446	9	Ades8385 Human gen
19	15.8	79.0	403	8	Ach48692 Human leu
20	15.8	79.0	467	8	Ach40016 Human foe
21	15.8	79.0	712	7	ACD95832 Human col
22	15.8	79.0	810	5	Abal6524 Human ner
23	15.8	79.0	1152	2	Aax14434 H. pylori

C 24	15.8	79.0	1255	9	Ades56623 Human gen
C 25	15.8	79.0	1255	9	Ades56627 Human gen
C 26	15.8	79.0	1255	9	Ades56615 Human gen
C 27	15.8	79.0	1255	9	Ades56619 Human gen
C 28	15.8	79.0	1774	2	Aaz42126 Human nor
C 29	15.8	79.0	3153	2	AAV06251 Human pur
C 30	15.8	79.0	3362	2	AAV06278 Human pur
C 31	15.8	79.0	3747	2	AAV06250 Human pur
C 32	15.8	79.0	4049	4	AAH02878 Human she
C 33	15.8	79.0	4170	4	AAH02878 Human she
C 34	15.8	79.0	4170	10	AAH02878 Human she
C 35	15.8	79.0	24053	4	Ades77056 Human pro
C 36	15.8	79.0	32768	2	Ades77056 Human pro
C 37	15.4	77.0	673	6	Ades77056 Human pro
C 38	15.4	77.0	673	9	Ades77056 Human pro
C 39	15.4	77.0	2067	9	Ades77056 Human pro
C 40	15.4	77.0	2067	6	Ades77056 Human pro
C 41	15.4	77.0	3685	6	Ades77056 Human pro
C 42	15.4	77.0	3685	7	Ades77056 Human pro
C 43	15.4	77.0	4112	5	Ades77056 Human pro
C 44	15.4	77.0	4137	3	Ades77056 Human pro
C 45	15.4	77.0	6029	4	Ades77056 Human pro

## ALIGNMENTS

RESULT 1  
ABK61505  
ID ABK61505 standard; DNA; 20 BP.  
XX  
AC ABK61505;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human NOV13 RT-PCR primer #1.  
XX  
KW Human; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;  
cell signal processing disorder; metabolic pathway modulation disorder;  
diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; primer;  
uterus cancer; immune response; graft-versus-host disease; Exon linking;  
acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;  
hypertension; congenital heart defects; multiple sclerosis; inflammation;  
Albright hereditary osteodystrophy; reverse transcriptase PCR.  
XX  
OS Homo sapiens.  
XX  
PN WO200216599-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 27-AUG-2001; 2001WO-US026510.  
XX  
PR 25-AUG-2000; 2000US-0228191P.  
PR 08-FEB-2001; 2001US-0267300P.  
PR 20-FEB-2001; 2001US-0269961P.  
PR 20-MAR-2001; 2001US-0277337P.  
XX  
(CURA-) CURAGEN CORP.  
PA (CORT-) COR THERAPEUTICS INC.  
XX  
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;  
Spytek KA, Szekeres ES, Tomlinson JB, Topper JN, Yang R;  
XX  
DR WPI; 2002-280937/32.  
PT New polypeptides for treating or preventing a disorder associated with  
them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.  
XX  
PS Example 2; Page 234; 263pp; English.  
CC The invention relates to an isolated polypeptide (NOVX) a mature form of  
NOVX, a NOVX variant (differing by no more than 15%), the nucleotide

CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
 CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it  
 CC and antibody against it, are useful for treating or preventing (e.g. by  
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOVX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation or Albrit  
 CC hereditary osteodystrophy and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC for developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence is a reverse transcriptase (RT)-PCR primer used to measure  
 CC tissue specific expression of mRNA encoding a NOVX protein

XX  
 SQ Sequence 20 BP; 3 A; 3 C; 8 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGACAGGTAGGGCTTTG 20  
 |||||  
 Db 1 CTGGACAGGTAGGGCTTTG 20

RESULT 2  
 ID ABQ98669 standard; DNA; 444 BP.  
 AC ABQ98669;  
 XX  
 DT 04-NOV-2002 (first entry)  
 XX  
 DE Human ORF476 coding sequence.  
 XX  
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulneryary;  
 KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2002082206-A1.  
 XX  
 PD 27-JUN-2002.  
 XX  
 PF 30-MAY-2001; 2001US-00867550.  
 XX  
 PR 30-MAY-2000; 2000US-0208427P.  
 XX  
 XX (LEAC/) LEACH M. D.  
 PA (MEHR/) MEHRABAN F.  
 PA (CONL/) CONLEY P. B.  
 PA (TOPP/) TOPPER J. N.  
 PA (LAWD/) LAW D.

XX Leach WD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX  
 DR WPI; 2002-626554/67.  
 DR P-PSDB; ABP64106.

XX New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including

PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 XX  
 PS Claim 2; SEQ ID NO 951; 78pp; English.  
 XX  
 CC The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?DocID=20020082206

XX SQ Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGGACAGGTAGGGCTTTG 20  
 |||||  
 Db 60 CTGGACAGGTAGGGCTTTG 41

RESULT 3  
 AAS74747  
 ID AAS74747 standard; cDNA; 445 BP.  
 XX  
 AC AAS74747;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #10551.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Drmanac RT, Liu C, Tang YT;  
 XX  
 XX WPI; 2001-639362/73.  
 DR P-PSDB; ABG10360.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 1; SEQ ID NO 10551; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activities. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAG64197-AA94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 445 BP; 89 A; 112 C; 143 G; 101 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 5; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20  
 |||||  
 Db 168 CTGGACAGGTTAGGCTTTG 187

RESULT 4  
 ABQ98670/c

ID ABQ98670 standard; DNA; 763 BP.

XX AC ABQ98670;

XX DT 04-NOV-2002 (first entry)

XX DE Human ORF477 coding sequence.

XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Valnerary;  
 KW Antiflamatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.

XX OS Homo sapiens.

XX PN US2002082206-A1.

XX PD 27-JUN-2002.

XX PF 30-MAY-2001; 2001US-00867550.

XX PR 30-MAY-2000; 2000US-0208427P.

XX PA (LEAC/) LEACH M D.

XX PA (MEHR/) MEHRABAN F.

XX PA (CONL/) CONLEY P B.

XX PA (TOPP/) TOPPER J N.

XX PA (LAWD/) LAW D.

XX PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

XX DR WPI; 2002-626554/67.

XX DR P-PSDB; ABP64107.

XX PT New polypeptide designated ORFX are present in human atherogenic cells  
 CC and are useful to prevent and treat ORFX-associated disorders including  
 CC cancer, allergy, wound healing or autoimmune, cardiovascular or  
 CC inflammatory disease.

XX PS Claim 2; SEQ ID NO 953; 789pp; English.

XX CC The present invention relates to novel human ORFX polypeptides and their

CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?docID=20020082206

XX SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 763;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20

|||||

Db 189 CTGGACAGGTTAGGCTTTG 170

RESULT 5

ABK61465/c

ID ABK61465 standard; cDNA; 1183 BP.

XX AC ABK61465;

XX DT 18-JUN-2002 (first entry)

XX DE Human cDNA encoding protein NOV13.

XX KW Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;  
 KW cell signal processing disorder; metabolic pathway modulation disorder;  
 KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;  
 KW uterus cancer; immune response; graft-versus-host disease;  
 KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;  
 KW hypertension; congenital heart defects; multiple sclerosis; inflammation;  
 KW Albright hereditary osteodystrophy.

XX OS Homo sapiens.

XX PN WO200216599-A2.

XX PD 28-FEB-2002.

XX PF 27-AUG-2001; 2001WO-US026510.

XX PR 25-AUG-2000; 2000US-0228191P.

XX PR 08-FEB-2001; 2001US-0267300P.

XX PR 20-FEB-2001; 2001US-0269961P.

XX PR 20-MAR-2001; 2001US-0277337P.

XX PA (CURA-) CURAGEN CORP.

XX PA (COR-) COR THERAPEUTICS INC.

XX PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;

XX PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;

XX DR WPI; 2002-280937/32.

XX DR P-PSDB; AAU91308.

XX PT New polypeptides for treating or preventing a disorder associated with  
 CC them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.

XX PS Claim 1; Page 98; 263pp; English.

XX CC The invention relates to an isolated polypeptide (NOVX) a mature form of  
 CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide  
 CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
 CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it

CC and antibody against it, are useful for treating or preventing (e.g. by  
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOVX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation or Albrit  
 CC hereditary osteodystrophy and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC for developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence encodes a NOVX protein  
 XX  
 SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGACAGGTAGGGCTTTG 20  
 DB 301 CTGGACAGGTAGGGCTTTG 282

RESULT 6  
 AAD43980/c  
 ID AAD43980 standard; cDNA; 2567 BP.  
 XX  
 AC AAD43980;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 DE Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.

XX Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective; gene; ss.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH 415..1200  
 FT CDS /\*tag= a  
 FT /\*product= "Human SLAP-2"  
 XX  
 PN WO200242457-A1.  
 XX  
 PD 30-MAY-2002.  
 XX  
 PF 20-NOV-2001; 2001WO-US043367.  
 XX  
 PR 22-NOV-2000; 2000US-0252545P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;  
 XX WPI; 2002-463632/49.  
 DR P-PSDB; AAE26357.  
 XX

PT Novel substantially purified human SH2/SH3-domain-containing adapter  
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic

PT intervention in immunological and inflammatory disorders and cancer.  
 XX  
 PS Claim 2; Fig 1; 85pp; English.  
 XX  
 CC The invention relates to a substantially purified human SH2/SH3-domain-  
 CC containing adapter polypeptide, termed Src-like Adapter Protein-2 (SLAP-  
 CC 2). The invention is useful for treating an immune disorder involving  
 CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is  
 CC useful for screening for antagonists or inhibitors of the interaction of  
 CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or  
 CC preventing diseases or disorders associated with aberrant or uncontrolled  
 CC cellular signal transduction, for determining those cellular signalling  
 CC molecules which associate with hSLAP-2 and which provide critical signals  
 CC for cell activation, and as effectors in methods to affect i- cell  
 CC activation. The invention is useful in screening assays to identify and  
 CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for  
 CC potential use to treat autoimmune diseases which may be caused by  
 CC hyperactivated B cells, as well as to treat diseases which may be caused  
 CC by hyperactivated T cells, in addition to other immune system related  
 CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,  
 CC inflammation disorders, diseases and conditions, rheumatoid arthritis,  
 CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's  
 CC and ulcerative colitis), allergies, particularly those involving  
 CC hyperactivity of B-cells and T- cells, or other immune cells, such as  
 CC mast cells or eosinophils, autoimmune diseases such as systemic lupus  
 CC erythematosus and multiple sclerosis, pulmonary diseases including  
 CC asthma, acute respiratory distress syndrome, and chronic obstructive  
 CC pulmonary disorder, tissue/ organ rejection and cancer. The invention is  
 CC useful in gene therapy. The present sequence is human SLAP-2 cDNA  
 XX  
 SQ Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 2567;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGACAGGTAGGGCTTTG 20  
 DB 318 CTGGACAGGTAGGGCTTTG 299

RESULT 7  
 ADD22694/c  
 ID ADD22694 standard; DNA; 538 BP.  
 XX  
 AC ADD22694;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Filamentous fungi Aspergillus genus DNA sequence, SEQ ID NO 145.

XX promoter activity; filamentous fungi Aspergillus genus;  
 KW transcription activation; glucose; gene transcription; ds.  
 XX  
 OS Aspergillus oryzae.  
 XX  
 PN JP2003144171-A.  
 XX  
 PD 20-MAY-2003.  
 XX  
 PF 16-NOV-2001; 2001JP-00351368.  
 XX  
 PR 16-NOV-2001; 2001JP-00351368.  
 XX  
 PA (HGET ) HIGETA SHOYU KK.  
 PA (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.  
 XX  
 DR WPI; 2003-818168/77.  
 XX  
 PT Novel promoter DNA derived from Aspergillus genus, useful for protein  
 PT production.  
 XX  
 PS Claim 1; SEQ ID NO 145; 83pp; Japanese.

XX CC The invention relates to a novel DNA sequence comprising any one of 74  
CC sequences with promoter activity, having a fully defined sequence shown  
CC in the specification or having a sequence which hybridised under  
CC stringent conditions to one of the 74 sequences. The invention further  
CC comprises one of the 74 DNA sequences or their DNA fragments, derived  
CC from filamentous fungi *Aspergillus* genus, being useful for manufacturing  
CC a protein. One of the DNA sequences or its fragment can effectively  
CC activate transcription of a gene encoding a target protein irrespective  
CC of the presence or absence of glucose. The DNA sequences and their  
CC fragments can constantly activate gene transcription. This polynucleotide  
CC sequence represents one of the 74 filamentous fungi *Aspergillus* genus DNA  
CC sequences of the invention.

XX SQ Sequence 538 BP; 128 A; 123 C; 132 G; 155 T; 0 U; 0 Other;  
Query Match 84.0%; Score 16.8; DB 9; Length 538;  
Best Local Similarity 90.0%; Pred. No. 1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CTGGACAGGTTAGGGCTTTG 20  
|||||  
DB 494 CTGGACAGGTTAGGGCTTTG 475

RESULT 8  
ABZ56309/c  
ID ABZ56309 standard; cDNA; 589 BP.  
XX AC ABZ56309;  
XX DT 28-MAR-2003 (first entry)  
XX DE *Aspergillus oryzae* polynucleotide SEQ ID NO 5422.  
XX KW *Aspergillus oryzae*; fermentation; fungus; industrial; EST;  
XX KW expressed sequence tag; gene; ss.  
XX OS *Aspergillus oryzae*.  
XX XX WO200279476-A1.  
XX PD 10-OCT-2002.  
XX PF 22-MAR-2002; 2002WO-IB000890.  
XX PR 30-MAR-2001; 2001JP-00098371.  
XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX PA (NARE-) NAT RES INST BREWING.  
XX PA (NORQ) NAT FOOD RES INST MIN AGRIC.  
XX PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;  
XX PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;  
XX WPI; 2003-046817/04.  
XX PT Detection of expression of specific *Aspergillus* genes for monitoring the  
XX fermentation and growth conditions of the fungus, using DNA probes.  
XX PS Claim 1; SEQ ID NO 5422; 48pp + Sequence Listing; Japanese.  
XX CC The invention relates to a polynucleotide having any of 6006 specific  
CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under  
CC specific culture conditions including one or more of eutrophic,  
CC oligotrophic, solid, early germination, alkaline, high temperature, low  
CC temperature or maltose culture or polynucleotides stringently hybridising  
CC to these sequences. The polynucleotides are useful for monitoring the  
CC progress of fermentation and the growth conditions of a fungus,  
CC especially of *Aspergillus oryzae* which is widely used in industrial  
CC fermentation. Also monitoring for fungal contamination. Note: the  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 589 BP; 143 A; 210 C; 95 G; 141 T; 0 U; 0 Other;  
Query Match 84.0%; Score 16.8; DB 7; Length 589;  
Best Local Similarity 90.0%; Pred. No. 1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CTGGACAGGTTAGGGCTTTG 20  
|||||  
DB 100 CTGGACAGGTTAGGGCTTTG 81

RESULT 9  
AAF14531/c  
ID AAF14531 standard; cDNA; 809 BP.  
XX AC AAF14531;  
XX DT 13-MAR-2001 (first entry)  
XX DE *Aspergillus oryzae* EST SEQ ID NO:7054.  
XX KW Multiple gene expression; filamentous fungal cell; EST;  
XX KW expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;  
XX KW *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombination;  
XX KW culture condition; environmental stress; spore morphogenesis;  
XX KW metabolic pathway engineering; catabolic pathway engineering; ss.  
XX OS *Aspergillus oryzae*.  
XX XX WO200056762-A2.  
XX PD 28-SEP-2000.  
XX PF 22-MAR-2000; 2000WO-US007781.  
XX PR 22-MAR-1999; 99US-00273623.  
XX PA (NOVO) NOVO NORDISK BIOTECH INC.  
XX PA (NOVO) NOVO NORDISK AS.  
XX PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
XX WPI; 2000-594572/56.  
XX PT Monitoring differential expression of genes in filamentous fungal cells  
XX PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
XX PT substrate of expressed sequence tags.  
XX PS Claim 88; Page 2863-2864; 3161pp; English.  
XX CC The present invention describes a method for monitoring differential  
XX expression of genes in a first filamentous fungal (FF) cell relative to  
XX expression of the same genes in one or more second filamentous fungal  
XX cells. The method uses fluorescence-labeled nucleic acids isolated from  
XX the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
XX are used in the methods for monitoring differential expression of genes  
XX in a first filamentous fungal (FF) cell relative to expression of the  
XX same genes in one or more second filamentous fungal cells. Monitoring the  
XX global expression of genes from FF cells allows the production potential  
XX of the microorganism to be improved. New genes may be discovered,  
XX possible functions of unknown open reading frames can be identified and  
XX gene copy number variation and stability can be monitored. The expression  
XX of genes can be used to study how FF cells adapt to changes in culture  
XX conditions, environmental stress, spore morphogenesis, recombination,  
XX metabolic or catabolic pathway engineering. Using ESTs provides several  
XX advantages over genomic or random cDNA clones including elimination of  
XX redundancy as one spot on an array equals one gene or open reading frame,  
XX and organisation of the microarrays based on function of the gene  
XX products to facilitate analysis of the results. AAF07478 to AAF11247  
XX represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents  
XX ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from

CC Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from  
 CC Trichoderma reesei, which are all specifically claimed in the present  
 CC invention

SQ Sequence 809 BP; 182 A; 265 C; 161 G; 199 T; 0 U; 2 Other;  
 Query Match 84.0%; Score 16.8; DB 3; Length 809;  
 Best Local Similarity 90.0%; Pred. No. 1.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20  
 |||||  
 Db 62 CTGGACAGGTTAGGCTTTG 43

RESULT 10  
 ADA53227/c  
 ID ADA53227 standard; cDNA; 2305 BP.

XX ADA53227;  
 AC  
 DT 20-NOV-2003 (first entry)  
 XX Human coding sequence, SEQ ID 795.

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
 KW inflammatory disease; osteoporosis; neurological disease; gene; ss.  
 XX  
 OS Homo sapiens.

XX EP1293569-A2.  
 PN  
 XX 19-MAR-2003.

XX 21-MAR-2002; 2002EP-00006586.

XX 14-SEP-2001; 2001JP-00328381.

XX 24-JAN-2002; 2002US-0350435P.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-395539/38.

DR P-PSDB; ADA54866.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 1; SEQ ID NO 795; 205pp; English.

XX The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 2305 BP; 457 A; 730 C; 621 G; 497 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 7; Length 2305;  
 Best Local Similarity 90.0%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20  
 |||||  
 Db 2118 CTGGACAGGTTAGGCTGTG 2099

RESULT 11  
 ADA02960/c  
 ID ADA02960 standard; DNA; 44325 BP.

XX ADA02960;

XX AC

XX 06-NOV-2003 (first entry)

XX Human DAD1 carcinoma associated gene, SEQ ID NO:1478.

XX Human; carcinoma associated; oncogene; carcinoma; cancer; breast;  
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
 KW gene; ds.

XX OS Homo sapiens.

XX WO2003057146-A2.

XX 17-JUL-2003.

XX 26-DEC-2002; 2002WO-US041414.

XX 26-DEC-2001; 2001US-00035832.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-597068/55.

XX New recombinant nucleic acid encoding carcinoma associated protein,  
 PT useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 1478; 245pp; English.

XX The invention relates to recombinant carcinoma associated (CA) nucleic  
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to  
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The  
 CC invention also encompasses expression vectors and host cells comprising a  
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically  
 CC binds to the protein, and a biochip comprising CA nucleic acid or  
 CC fragments thereof. The sequences of the invention were identified using  
 CC oncogenic retroviruses, which insert into the genome of the host organism  
 CC at random. Many of these do not carry transduced host oncogenes or  
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a  
 CC direct consequence of the effects of proviral integration into host  
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
 CC leukaemia) or a propensity to carcinoma by determination of the sequence  
 CC of a CA gene, or by determination of CA gene expression in particular  
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
 CC therapeutic agents and in screening and evaluating drug candidates. The  
 CC present sequence represents a specifically claimed human CA nucleic acid  
 CC segment of the invention. Note: The complete sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 44325 BP; 11218 A; 9675 C; 9902 G; 13530 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 8; Length 44325;  
 Best Local Similarity 90.0%; Pred. No. 1.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20

Db 38075 CTGGACAGGTTAGGCTGTG 38056

RESULT 12

ID ADB72698/c

AD72698 standard; DNA; 44325 BP.

XX

AC ADB72698;  
 XX 04-DEC-2003 (first entry)  
 XX Human DAD1 gene.  
 DE human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;  
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.  
 XX Homo sapiens.  
 OS WO2003008583-A2.  
 XX 30-JAN-2003.  
 PD 26-DEC-2001; 2001WO-US051291.  
 XX 02-MAR-2001; 2001US-00798596.  
 XX 23-OCT-2001; 2001US-00004113.  
 PR 08-NOV-2001; 2001US-00052482.  
 PR 30-NOV-2001; 2001US-00397722.  
 PR 20-DEC-2001; 2001US-00034650.  
 XX (SAGR-) SAGRES DISCOVERY.  
 PA Morris DW, Engelhard EK;  
 XX WPI; 2003-239337/23.  
 DR New recombinant nucleic acid, useful for treating carcinomas, lymphomas,  
 XX cancers, neoplasm, adenocarcinoma, or sarcomas.  
 PT Claim 1; SEQ ID NO 526; 2304pp; English.  
 XX The invention relates to a novel recombinant nucleic acid comprising a  
 CC nucleotide sequence selected from any of the 660 sequences fully defined  
 CC in the specification. A polynucleotide of the invention has cytostatic  
 CC activity, and may have a use in gene therapy, or in a vaccine. The  
 CC recombinant nucleic acids and polypeptides are useful for treating  
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and  
 CC sarcomas. The present sequence represents a human gene of the invention.  
 XX Sequence 44325 BP; 11218 A; 9675 C; 9902 G; 13530 T; 0 U; 0 Other;  
 SQ Query Match 84.0%; Score 16.8; DB 9; Length 44325;  
 Best Local Similarity 90.0%; Pred. No. 1.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CTGGACAGGTTAGGGCTTTG 20  
 DB 38075 CTGGACAGGATAGGGCTGTG 38056  
 RESULT 13  
 ID ADC85440/c  
 XX ADC85440 standard; DNA; 44325 BP.  
 AC ADC85440;  
 XX 01-JAN-2004 (first entry)  
 DT Human Dad1 genomic sequence.  
 DE Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;  
 KW secreted; transmembrane; intracellular; ds.  
 XX Homo sapiens.  
 OS WO2003045230-A2.  
 XX 05-JUN-2003.  
 PD 02-DEC-2002; 2002WO-US038582.  
 PP

XX 30-NOV-2001; 2001US-00997722.  
 PR (SAGR-) SAGRES DISCOVERY.  
 PA Morris DW, Engelhard EK;  
 XX WPI; 2003-513603/48.  
 DR New recombinant nucleic acid comprising a nucleotide sequence of any of  
 PT the carcinoma-associated (CA) genes, useful for screening for drug  
 PT candidates for diagnosing or treating carcinomas.  
 XX Claim 1; SEQ ID NO 226; 983pp; English.  
 XX The invention relates to a recombinant nucleic acid comprising a  
 CC nucleotide sequence selected from any of the fully defined carcinoma-  
 CC associated (CA) genes from the 50 tables given in the specification. The  
 CC CA proteins are secreted, transmembrane or intracellular proteins. The  
 CC recombinant nucleic acids are useful for screening for drug candidates  
 CC for diagnosing or treating carcinomas. Sequences given in ADC85215-  
 CC ADC85514 represent CA genes of the invention.  
 XX Sequence 44325 BP; 11218 A; 9675 C; 9902 G; 13530 T; 0 U; 0 Other;  
 SQ Query Match 84.0%; Score 16.8; DB 9; Length 44325;  
 Best Local Similarity 90.0%; Pred. No. 1.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CTGGACAGGTTAGGGCTTTG 20  
 DB 38075 CTGGACAGGATAGGGCTGTG 38056  
 RESULT 14  
 ADE57896  
 ID ADE57896 standard; DNA; 2446 BP.  
 XX ADE57896;  
 AC 29-JAN-2004 (first entry)  
 DT Human gene U72649, SEQ ID NO 3762.  
 XX Human; ds; gene; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX Homo sapiens.  
 OS WO2003016475-A2.  
 XX 27-FEB-2003.  
 PD 14-AUG-2002; 2002WO-US025765.  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX (GEHC) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 DR GENBANK; U72649.  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX Claim 1; Page; 1017pp; English.  
 XX

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human DNA (shown in Table 2 of the  
CC specification) which encodes one of the polypeptides of the invention  
CC which is differentially expressed during pain. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2446 BP; 567 A; 637 C; 596 G; 646 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 9; Length 2446;  
Best Local Similarity 94.4%; Pred. NO. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 GGACAGGTTAGGGCTTTG 20  
|||||  
Db 1850 GGACAGGTTAGGGCTTTG 1867

RESULT 15

AD5E7892

ID AD5E7892 standard; DNA; 2446 BP.

AC AD5E7892;

DT 29-JAN-2004 (first entry)

DE Human gene U72649, SEQ ID NO 3758.

KW Human; ds; gene; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.

OS Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; U72649.

XX

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human DNA (shown in Table 2 of the  
CC specification) which encodes one of the polypeptides of the invention  
CC which is differentially expressed during pain. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 2446 BP; 567 A; 637 C; 596 G; 646 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 9; Length 2446;  
Best Local Similarity 94.4%; Pred. NO. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 GGACAGGTTAGGGCTTTG 20

Db 1850 GGACAGGTTAGGGCTTTG 1867

Search completed: February 19, 2004, 21:51:42  
Job time : 21.5292 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:54 ; Search time 132.422 Seconds  
(without alignments)  
4510.152 Million cell updates/sec

Title: US-09-939-853A-140  
Perfect score: 20  
Sequence: 1 ctggacaggttagggcttgg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hct.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hct.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_pbg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	616	13	BX383606
C 2	20	100.0	778	12	BG178487
C 3	20	100.0	878	12	BQ053486
C 4	20	100.0	986	12	BQ054265

C 5	20	100.0	1020	12	BQ054281
C 6	20	100.0	1201	9	AL541041
C 7	18.4	92.0	901	13	BU327205
C 8	18.4	92.0	962	14	CA987486
C 9	17.4	87.0	332	9	AA719167
C 10	17.4	87.0	357	9	AA759254
C 11	17.4	87.0	413	13	BX089099
C 12	17.4	87.0	740	29	AG048198
C 13	16.8	84.0	177	29	CG591538
C 14	16.8	84.0	212	14	CB487227
C 15	16.8	84.0	259	10	BB306076
C 16	16.8	84.0	316	14	CB486524
C 17	16.8	84.0	380	29	AG214612
C 18	16.8	84.0	415	10	BE665847
C 19	16.8	84.0	470	10	BF448079
C 20	16.8	84.0	498	28	BH258504
C 21	16.8	84.0	545	12	BI529969
C 22	16.8	84.0	559	13	BU648008
C 23	16.8	84.0	624	29	CE663064
C 24	16.8	84.0	628	14	CB445363
C 25	16.8	84.0	645	12	BM003795
C 26	16.8	84.0	695	29	CE509914
C 27	16.8	84.0	721	14	CF206129
C 28	16.8	84.0	727	10	BF215517
C 29	16.8	84.0	749	29	AG053044
C 30	16.8	84.0	757	28	BH012585
C 31	16.8	84.0	907	10	BF681379
C 32	16.8	84.0	942	10	BF310891
C 33	16.8	84.0	989	10	BE729761
C 34	16.8	84.0	990	10	BF699088
C 35	16.4	82.0	206	29	CG621536
C 36	16.4	82.0	223	9	AU254349
C 37	16.4	82.0	283	10	AW799548
C 38	16.4	82.0	297	29	CE360042
C 39	16.4	82.0	363	14	CD995897
C 40	16.4	82.0	368	28	AQ423959
C 41	16.4	82.0	378	9	AV208185
C 42	16.4	82.0	423	10	AW558403
C 43	16.4	82.0	541	28	AZ700347
C 44	16.4	82.0	577	29	CNS07H2B
C 45	16.4	82.0	578	28	AZ850663

## ALIGNMENTS

RESULT 1  
BX383606/c  
LOCUS BX383606 Homo sapiens T CELLS (TUKAT CELL LINE) COT 10-NORMALIZED  
DEFINITION Homo sapiens cDNA clone CS0DU013YK10 5-PRIME, mRNA sequence.  
ACCESSION BX383606  
VERSION BX383606.1 GI:30457152  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
I. (bases 1 to 616)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9825.r For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DU013BF05QPI&cluster=9825.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ invitrogen Corporation 1500

Paraday Avenue Genoscope sequence ID : CS0DJ013BF05QPL.

## FEATURES

source

1. .616  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DJ013VK10"  
 /cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
 /cell\_line="JURKAT"  
 /clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 616;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20

Db 350 CTGGACAGGTTAGGGCTTTG 331

## RESULT 2

BG178487/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG178487 778 bp mRNA linear EST 06-FEB-2001  
 602328305F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4429896 5',  
 mRNA sequence.

BG178487

BG178487.1 GI:12685190

EST.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 778)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAW10182 row: 1 column: 01  
 High quality sequence stop: 657.  
 Location/Qualifiers  
 1. .778  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4429896"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 91"  
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

## FEATURES

source

1. .616  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4429896"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 91"  
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 778;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20

Db 265 CTGGACAGGTTAGGGCTTTG 246

## RESULT 3

BQ053486/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ053486 878 bp mRNA linear EST 29-MAR-2002  
 AGENCOURT\_6820248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5935253  
 5', mRNA sequence.

BQ053486

BQ053486.1 GI:19812826

EST.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 878)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1CM2122 row: 1 column: 06  
 High quality sequence stop: 394.  
 Location/Qualifiers  
 1. .878  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5935253"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## FEATURES

source

1. .878  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5935253"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 878;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20

Db 176 CTGGACAGGTTAGGGCTTTG 157

## RESULT 4

BQ054265/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

BQ054265 986 bp mRNA linear EST 29-MAR-2002  
 AGENCOURT\_6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
 5', mRNA sequence.

BQ054265

BQ054265.1 GI:19813605

EST.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

1 (bases 1 to 986)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1CM2122 row: 1 column: 06  
 High quality sequence stop: 394.  
 Location/Qualifiers  
 1. .878  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5935253"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LICM2125 row: i column: 12  
 High quality sequence stop: 515.

FEATURES source  
 1..986  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5936339"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 986;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20  
 |||||  
 Db 182 CTGGACAGGTTAGGCTTTG 163

RESULT 5  
 BQ054281/c  
 LOCUS BQ054281 1020 bp mRNA linear EST 29-MAR-2002  
 DEFINITION AGENCOURT 6830234 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936362  
 5', mRNA sequence.  
 ACCESSION BQ054281  
 VERSION BQ054281.1 GI:19813621  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1020)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LICM2125 row: j column: 11  
 High quality sequence stop: 556.

FEATURES source  
 1..1020  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"

/db\_xref="taxon:9606"  
 /clone="IMAGE:5936362"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 1020;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20  
 |||||  
 Db 115 CTGGACAGGTTAGGCTTTG 96

## RESULT 6

AL541041/c  
 LOCUS AL541041 1201 bp mRNA linear EST 12-MAY-2003  
 DEFINITION AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YK23  
 5-PRIME, mRNA sequence.  
 ACCESSION AL541041  
 VERSION AL541041.2 GI:30544829  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1201)  
 AUTHORS Li W.B., Gruber C., Jessee J. and Palayes D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12871733.  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [secre@genoscope.cns.fr](mailto:secre@genoscope.cns.fr), Web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9825.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DE005AF12QP1&cluster=9825.r>. Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DE005AF12QP1.

## FEATURES source

1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE005YK23"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1201;  
 Best Local Similarity 100.0%; Pred. No. 98;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20

```

Db          362 CTGGACAGGTTAGGCGTTTG 343
|||||
RESULT 7
BU327205
LOCUS
DEFINITION BU327205 901 bp mRNA linear EST 28-NOV-2002
sequence.
ACCESSION BU327205
VERSION BU327205.1 GI:25835206
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 901)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..901
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST392m20"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCIN63"
/note="Organ: heads; Vector: pBluescript II KS(+); Site: 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methyated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Query Match 92.0%; Score 18.4; DB 13; Length 901;
Best Local Similarity 95.0%; Pred. No. Se+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20
|||||
Db 659 CTGGACGCTTAGGCGTTTG 678
|||||

RESULT 8
CA987486/c
LOCUS
DEFINITION CA987486 962 bp mRNA linear EST 06-JAN-2003
AGENCOURT 11284551 NICHG XGC Emb1 Xenopus laevis cDNA clone
IMAGE:6864145 5', mRNA sequence.
ACCESSION CA987486

```

---

```

VERSION CA987486.1 GI:27520161
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 962)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14484 row: g column: 24
High quality sequence start: 228
High quality sequence stop: 380.
Location/Qualifiers
1..962
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6864145"
/tissue_type="embryo (stage 10)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG XGC Emb1"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 1.55 kb. Constructed by Life Technologies. Note: This
is a Xenopus Gene Collection (XGC) library."

ORIGIN
Query Match 92.0%; Score 18.4; DB 14; Length 962;
Best Local Similarity 95.0%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20
|||||
Db 840 CTGGCAGGTTAGGCGTTTG 821
|||||

RESULT 9
AA719167/c
LOCUS
DEFINITION AA719167 332 bp mRNA linear EST 12-JAN-1999
ah47e08.s1 Soares_testis_NHT Homo sapiens cDNA clone 1292678 3',
mRNA sequence.
ACCESSION AA719167
VERSION AA719167.1 GI:2732266
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 332)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

```

www-bio.llnl.gov/bbrp/image/image.html  
 Insert length: 472 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 316.  
 Location/Qualifiers

# FEATURES

source

1..332

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="1292678"  
 /sex="male"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares testis NHT"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech  
 Laboratories, Inc., and primed with a Not I - oligo(dT)  
 primer [5].  
 TGTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'.  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization to Cots5, and was  
 constructed by Bento Soares and M. Fatima Bernaldo."

## ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 332;  
 Best Local Similarity 94.7%; Pred. No. 1.2e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTAGGGCTTT 19

Db 207 CTGGACAGGTAGGGCTTT 189

## RESULT 10

AA759254/c  
 LOCUS  
 DEFINITION  
 aa49254.sl Soares\_testis\_NHT Homo sapiens cDNA clone 1322374 3',  
 mRNA sequence.  
 AA759254.1 GI:2807117  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 357)  
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bernaldo, Ph.D.

## JOURNAL

### COMMENT

cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CCGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Insert length: 466 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham.  
 Location/Qualifiers

## FEATURES

source

1..357

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="1322374"  
 /sex="male"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares testis NHT"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified

## ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 357;  
 Best Local Similarity 94.7%; Pred. No. 1.2e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTAGGGCTTT 19

Db 206 CTGGACAGGTAGGGCTTT 188

## RESULT 11

EX089099  
 LOCUS  
 DEFINITION  
 EX089099 Soares testis\_NHT Homo sapiens cDNA clone IMAGp998P23333  
 ! IMAGE:1322374, mRNA Sequence.  
 EX089099  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 413)  
 Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,  
 Radcliff, J., Schneider, D. and Korn, B.  
 Human Unigeneset - RZPD3  
 Unpublished (2003)  
 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZPD: IMAGp998P233333.  
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLIB No.972)  
 http://www.rzpd.de/cloneCards/cgi-  
 bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubneweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de

This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.

## FEATURES

source

1..413

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGp998P233333 ; IMAGE:1322374"  
 /sex="male"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares testis NHT"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech  
 Laboratories, Inc., and primed with a Not I - oligo(dT)  
 primer [5].  
 TGTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'.  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization to Cots5, and was

constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 413;  
 Best Local Similarity 94.7%; Pred. No. 1.3e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTT 19  
 |||||  
 Db 214 CTGGACAGGTCAGGGCTTT 232

RESULT 12  
 AG048198/c  
 LOCUS  
 DEFINITION Pan troglodytes DNA, clone: PTB-027P05.R, genomic survey sequence.  
 ACCESSION AG048198  
 VERSION AG048198.1 GI:16585090  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes

REFERENCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
 Totoki, Y., Watanabe, H. and Sakaki, Y.  
 BAC end sequences of Library PTB  
 2 (bases 1 to 740)  
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
 Totoki, Y., Watanabe, H. and Sakaki, Y.  
 Direct Submission  
 Submitted (02-AUG-2001) Asso Fujiyama, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail: chimbesgsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,  
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
 Clones are derived from the chimpanzee BAC library PTB This BAC end  
 was generated during the Rad process and may have higher chance of  
 clone tracking errors.  
 PRIMERS  
 Sequencing: M13rev

LIBRARY  
 Vector : pKS145  
 R.Site 1 : SacI  
 R.Site 2 : SacI

FEATURES  
 source  
 1..740  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="PTB-027P05.R"  
 /sex="male"  
 /cell\_type="lymphoblast"  
 /clone\_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 29; Length 740;  
 Best Local Similarity 94.7%; Pred. No. 1.4e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGGACAGGTTAGGGCTTT 20  
 |||||  
 Db 583 TGGACAGGTTAGGGCTTT 565

RESULT 13  
 CG591538/c  
 LOCUS  
 DEFINITION OST246195 Mus musculus 129Sv/Ev Mus musculus genomic clone  
 OST246195, genomic survey sequence.  
 ACCESSION CG591538  
 VERSION CG591538.1 GI:37399554

GSS.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 177)  
 Zambronicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J.,  
 Piggett, J., Beltrandelio, H., Buxton, E.C., Edwards, J., Finch, R.A.,  
 Friddle, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C.,  
 Key, B.W. Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D.,  
 Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,  
 Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,  
 Zhu, Q., Person, C. and Sands, A.T.  
 Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap  
 screen to identify potential targets for therapeutic intervention  
 Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
 Contact: Zambrowicz BP  
 OmniBank  
 Lexicon Genetics Incorporated  
 4000 Research Forest Drive, The Woodlands, TX 77381, USA  
 Email: materials@lexgen.com  
 Gene trap sequence tag generated by 3' RACE from mouse ES cells as  
 described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)  
 Class: Gene Trap.  
 Location/Qualifiers  
 1..177  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strains="129Sv/Ev"  
 /db\_xref="taxon:10090"  
 /clone="OST246195"  
 /cell\_type="embryonic stem cell"  
 /clone\_lib="Mus musculus 129Sv/Ev"

ORIGIN

Query Match 84.0%; Score 16.8; DB 29; Length 177;  
 Best Local Similarity 90.0%; Pred. No. 2.1e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20  
 |||||  
 Db 44 CTGGACAGGATAGGGCTTAG 25

RESULT 14  
 CB487227/c  
 LOCUS  
 DEFINITION ONYKt1001012 Oncorhynchus mykiss reproductive Oncorhynchus mykiss  
 cDNA, mRNA sequence.  
 ACCESSION CB487227  
 VERSION CB487227.1 GI:29298453  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM  
 Oncorhynchus mykiss (rainbow trout)  
 Oncorhynchus mykiss  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopercyiformes; Teleostei; Euteleostei;  
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 1 (bases 1 to 212)  
 GRASP Consortium, Davidson, W.S., Koop, B.F. and  
 http://web.uvic.ca/cbr/grasp.  
 A survey of Salmo salar transcripts from high complexity cDNA  
 libraries  
 Unpublished (2002)  
 Contact: Koop BF  
 Centre for Biomedical Research  
 University of Victoria  
 PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada  
 Tel: 250 472 4067  
 Fax: 250 472 4075  
 Email: Bkoop@uvic.ca  
 Centre for Biomedical Research, University of Victoria cDNA  
 preparation and sequencing: Roberto Alberto, Marianne  
 Beetz-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, BF Koop.